

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 165 Seconds  
(without alignments)  
424.264 Million cell updates/sec

Title: us-09-445-576A-7

Perfect score: 960

Sequence: 1 EPTQPKKIVNAKDVNT.....KMFDKRCQLPYICQFGIV 181

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	181	2	AAW94256 Human tet
2	960	100.0	181	8	ADO24666 Tetranect
3	960	100.0	182	7	ADD27692 Human adi
4	960	100.0	186	5	ABG80910 Human phr
5	960	100.0	197	2	AAW94262 HGFXTN123
6	960	100.0	202	2	AAR60521 Human tet
7	960	100.0	202	4	AAM24034 Human EST
8	960	100.0	202	7	ADE09166 Novel pro
9	960	100.0	202	7	ADE34332 Tetranect
10	960	100.0	202	8	ADN04116 Antipeori
11	960	100.0	228	2	AAW94261
12	957	99.7	202	6	ABR48483 Human Pla
13	955	99.5	180	7	ADD27211 Human adi
14	936	97.5	186	5	ABG80912 Human phr
15	936	97.5	188	5	ABG80908 Human FX-
16	932.5	97.1	182	1	AAP60098 Sequence
17	914	95.2	173	7	ADD27481 Human adi
18	889	92.6	181	8	ADO24678 Tetranect
19	886	92.3	167	7	ADD27117 Human adi
20	883	92.0	181	8	ADO24677 Tetranect
21	880	91.7	181	8	ADO24676 Tetranect
22	878	91.5	173	7	ADD26862 Human adi
23	875	91.1	167	7	ADD27582 Human adi
24	875	91.1	180	2	AAW94264 HGFXTN23
25	835	87.0	157	7	ADD27429 Human adi

#### ALIGNMENTS

##### RESULT 1

AAW94256

ID AAW94256 standard; protein; 181 AA.

AC AAW94256;

XX

DT 20-MAR-2003 (revised)

DT 26-APR-1999 (first entry)

XX

DE Human tetranectin amino acid sequence.

XX

KW Trimeric polypeptide; tetranectin trimerising structural element; TTSE;  
KW fusion protein; ligand binding structure; toxin; enzyme; cytokine;  
KW artificial antibody; pharmacokinetic; pharmacodynamic; gene therapy;  
KW transfection; imaging; tumour; human; tetranectin.

OS Homo sapiens.

XX

PN W09856906-A1.

XX

PD 17-DEC-1998.

XX

PF 11-JUN-1998; 98WO-DK000245.

XX

PR 11-JUN-1997; 97DK-00000685.

XX

(THOG/) THOGERSEN H C.

PA (ETZE/) ETZERODT M.

PA (HOLT/) HOLTET T L.

PA (GRAV/) GRAVERSEN N J H.

PA (KAST/) KASTRUP J S.

PA (NIEL/) NIELSEN B B.

PA (LARS/) LARSEN I K.

XX

PI Thogersen HC, Eterodt M, Holtet TL, Graversen NJH, Kastrup JS;

PI Nielsen BB, Larsen IK;

XX

DR WPI; 1999-080897/07.

XX

PT New monomer polypeptide constructs for diagnosis and therapy - comprise a tetranectin trimerising structural element covalently linked to at least one heterologous moiety for providing functional activity.

PT

PS Claim 10; Page 60; 110pp; English.

XX

CC The invention relates to the design of trimeric polypeptides using polypeptide structural elements derived from the tetranectin protein family. The trimeric polypeptides constructed as a monomer polypeptide

CC

26 777.5 81.0 169 7 ADE08161 Novel pro  
27 777 80.9 186 5 ABG80916 Mouse Pmt  
28 777 80.9 187 5 ABG80914 Mouse FX-  
29 773.5 80.6 169 8 ABM83784 Human dia  
30 747 77.8 152 2 AAW94265 HGFXTN3 f  
31 743 77.4 137 5 ABG80906 Human hnt  
32 743 77.4 137 8 ADO24667 Tetranect  
33 743 77.4 142 5 ABG80911 Human phr  
34 723 75.3 143 5 ABG80909 Human FX-  
35 719 74.9 142 5 ABG80913 Human phr  
36 674 70.2 137 8 ADO24672 Tetranect  
37 672 70.0 137 8 ADO24674 Tetranect  
38 666 69.4 137 8 ADO24675 Tetranect  
39 663 69.1 137 8 ADO24673 Tetranect  
40 621 64.7 114 7 ADD27101 Human adi  
41 618 64.4 143 5 ABG80915 Mouse FX-  
42 617 64.3 142 5 ABG80917 Mouse Pmt  
43 448 46.7 89 5 ABG80918 Human phr  
44 422.5 44.0 197 3 AAB28526 Protein e  
45 422.5 44.0 197 5 AAU82642 Human bre

CC construct comprise at least one tetranectin trimerising structural  
 CC element (TTSE) which is covalently linked to at least one heterologous  
 CC moiety, the TTSE being capable of forming a stable complex with 2 other  
 CC TTSEs, with the proviso that the heterologous moiety is different from  
 CC any of the fusion proteins CIIH6PXTN123, H6FYTN123, H6FYTN12, H6FYTN23  
 CC (AAW94261 to AAW94264). The TTSE can be used for the construction of  
 CC conjugates with heterologous moieties such as a ligand binding structure,  
 CC a toxin, a detectable label, an in situ activatable substance, an enzyme,  
 CC a radioactive moiety, a cytokine, a non-proteinaceous polymer, a photo  
 CC cross-linking agent, or a group facilitating conjugation of the monomer  
 CC polypeptide construct to a target. They can be used as vehicles for  
 CC assembling antibody fragments into oligomeric or multivalent entities for  
 CC generating chimeric artificial antibodies having preselected  
 CC pharmacokinetic and/or pharmacodynamic properties. The constructs can be  
 CC used for targeted gene therapy involving selective delivery of the  
 CC material for transfection or infection of the specific population of  
 CC cells. They can also be used for delivering a substance to a cell or  
 CC tissue or for delivering an imaging or toxin-conjugated antibody to a  
 CC tumour. They can also be used for prevention or treating a disease or for  
 CC diagnosis. The TTSE provides a stable structure which can act as a  
 CC vehicle for a wide variety of conjugates. The present sequence represents  
 CC a human tetranectin sequence from which the TTSE can be derived. (Updated  
 CC on 20-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 181 AA;

Query Match 100.0%; Score 960; DB 2; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-93;  
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTKVHMKC 60  
 DB 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTKVHMKC 60  
 QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120  
 DB 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120  
 QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGAANGKWFDRKCRDOLPYICQFGI 180  
 DB 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGAANGKWFDRKCRDOLPYICQFGI 180  
 QY 181 V 181  
 DB 181 V 181

RESULT 2  
 ADO24666  
 ID ADO24666 standard; protein; 181 AA.  
 XX  
 AC ADO24666;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE Tetranectin protein.

XX d; antirheumatic; antiarthritic; antipsoriatic; anti-inflammatory;  
 KW antimicrobial; cytostatic; neuroprotective; gene therapy;  
 KW trimeric polypeptide; monomer; cytokine; trimerizing domain;  
 KW rheumatoid arthritis; psoriasis; Crohn's disease; infection; malignancy;  
 KW neurodegenerat.  
 XX Unidentified.  
 OS  
 XX WO2004039841-A2.  
 PN  
 XX 13-MAY-2004.  
 PD  
 XX 29-OCT-2003; 2003WO-DK000735.  
 PF  
 XX 29-OCT-2002; 2002DK-00001634.  
 PR  
 XX 29-OCT-2002; 2002US-0421807P.

XX (BORE-) BOREAN PHARMA AS.  
 PA (OTTO/) OTTOW H K.  
 XX Holtet TL, Andersen MH;  
 PI WPI; 2004-376164/35.  
 DR  
 XX New trimeric binding polypeptides for trimeric cytokines, useful for  
 PT preparing a pharmaceutical composition for the treatment of pathologies  
 PT mediated by trimeric cytokines, e.g. rheumatoid arthritis, psoriasis or  
 PT Crohn's disease.  
 XX  
 PS Example 9; SEQ ID NO 96; 96pp; English.  
 XX  
 CC The invention relates to a novel trimeric polypeptide comprising three  
 CC monomers, each of the monomers comprising a specific binding member  
 CC capable of binding a trimeric cytokine, and each of the monomers  
 CC comprising a trimerizing domain. The trimeric polypeptide is useful for  
 CC preparing a pharmaceutical composition that may be used for treating a  
 CC pathology mediated by a trimeric cytokine, such as rheumatoid arthritis,  
 CC psoriasis, Crohn's disease, infections, malignancies or neurodegenerative  
 CC diseases. This sequence corresponds to a protein used to generate the  
 CC trimeric polypeptide of the invention.  
 XX  
 SQ Sequence 181 AA;

Query Match 100.0%; Score 960; DB 8; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-93;  
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTKVHMKC 60  
 DB 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTKVHMKC 60  
 QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120  
 DB 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120  
 QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGAANGKWFDRKCRDOLPYICQFGI 180  
 DB 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGAANGKWFDRKCRDOLPYICQFGI 180  
 QY 181 V 181  
 DB 181 V 181

RESULT 3  
 ADD27692  
 ID ADD27692 standard; protein; 182 AA.  
 XX  
 AC ADD27692;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Human adipocyte Selected Interacting Domain (SID) prey protein #1149.

XX Adipocyte; protein-protein interaction; protein complex;  
 KW bait-prey complex; Selected Interacting Domain; SID; drug screening;  
 KW drug discovery; metabolic disease; obesity; lipodystrophy;  
 KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;  
 KW adipogenesis modulation; gene therapy; human.  
 XX Homo sapiens.  
 OS  
 XX WO200290544-A2.  
 PN  
 XX 14-NOV-2002.  
 PD  
 XX 03-MAY-2002; 2002WO-EP006333.  
 PF  
 XX 04-MAY-2001; 2001US-0288885P.

```
XX PA (HYBR-) HYBRIGENICS.
XX PA (LYNX-) LYNX THERAPEUTICS INC.
XX PI Legrain P, Whiteside S, Mao J, Khrebtukova I, Luo S;
XX WPI; 2003-111975/10.
XX DR N-PSDB; ADD27691.
XX PT New complex of bait and prey between two polypeptides or polynucleotides
XX PT encoding the two polypeptides of adipocytes, useful for selecting a
XX PT modulating compound that inhibits or activates protein-protein
XX PT interactions.
XX PS Claim 6; SEQ ID NO 1149; 232pp; English.
XX CC The invention relates to a bait-prey complex between two adipocyte
XX CC polypeptides, or between two polynucleotides encoding adipocyte
XX CC polypeptides. The invention also relates to selected interacting Domain
XX CC (SID) polypeptides which interact with selected bait polypeptides;
XX CC polynucleotides encoding SID polypeptides; vectors comprising SID
XX CC polynucleotides; recombinant host cells comprising an adipocyte
XX CC polynucleotide or a SID-encoding vector; a method of selecting for a
XX CC compound which modulates interactions between adipocyte polypeptides;
XX CC adipocyte modulator compounds identified using the method; a
XX CC pharmaceutical composition comprising an adipocyte modulator, or a SID-
XX CC encoding vector or host cell; and a protein chip comprising adipocyte
XX CC bait polypeptides. The bait-prey complexes of the invention are useful
XX CC for selecting a compound that inhibits or activates protein-protein
XX CC interactions between adipocyte polypeptides. The modulatory compounds
XX CC identified can be used in the treatment of metabolic diseases such as
XX CC obesity, lipodystrophy and type 2 diabetes mellitus, and in the
XX CC modulation of adipogenesis. The present sequence represents a
XX CC specifically claimed adipocyte SID prey polypeptide of the invention.
XX SQ Sequence 182 AA;
Query Match 100.0%; Score 960; DB 7; Length 182;
Best Local Similarity 100.0%; Pred. No. 8.9e-93;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTDLAQEVALLEQQALQTVCLKGTVMKMC 60
DB 2 EPPTQPKKIVNAKDVVNTKMFELKSRDLTDLAQEVALLEQQALQTVCLKGTVMKMC 61
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 62 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 121
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 180
DB 122 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 181
QY 181 V 181
DB 182 V 182
RESULT 4
ABG80910
ID ABG80910 standard; protein; 186 AA.
XX AC ABG80910;
XX DT 29-NOV-2002 (first entry)
XX DE Human PHTN protein.
XX KW Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;
XX KW beta-strand; connecting segment; 14loop region; tetranectin;
XX KW ligand-binding specificity; human.
XX OS Homo sapiens.

XX PN WO200248189-A2.
XX PD 20-JUN-2002.
XX PF 13-DEC-2001; 2001WO-DK000825.
XX PR 13-DEC-2000; 2000DK-00001872.
XX PR 28-FEB-2001; 2001US-0272098P.
XX PA (BORE-) BOREAN PHARMA AS.
XX PI Etzerodt M, Holtet TL, Graversen NUH, Thogersen HC;
XX DR WPI; 2002-643278/69.
XX DR N-PSDB; ABS66592.
XX PT Protein comprising a variant of model C-type lectin-like domains (CTLD),
XX PT in which alpha helices, beta-strands, connecting segments are conserved
XX PT to maintain CTLD scaffold structure, while the loop region is altered.
XX PS Claim 16; Fig 10; 168pp; English.
XX CC The present invention relates to a new protein with scaffold structure of a
XX CC C-type lectin-like domains (CTLD). The invention comprises a variant of a
XX CC model CTLD where alpha-helices and beta-strands and connecting segments
XX CC are conserved such that scaffold structure of C-type lectin-like domains
XX CC (CTLD) is substantially maintained, while the 14loop region is altered by
XX CC amino acid substitution, deletion, insertion or their combination. The
XX CC invention is useful for preparing a library of nucleotide sequences
XX CC encoding related proteins by randomising part or all of the nucleic acid
XX CC sequence encoding the loop region of its CTLD. The artificial CTLD
XX CC protein products are preferable to antibody derivatives as each binding
XX CC site is a single structurally autonomous protein domain. When used as
XX CC components of compositions to be used for in vivo diagnostic or
XX CC therapeutic purposes, artificial CTLD protein products constructed on the
XX CC basis of human CTLDs are virtually identical to the corresponding natural
XX CC CTLD protein already present in the body and are therefore less
XX CC immunogenic to the patient. They also have a smaller size, and thus
XX CC provide tissue penetration and distribution, as well as shorter half life
XX CC in circulation. Since murine and human tetranectin are identical in
XX CC structure, straightforward swapping of polypeptide segments defining
XX CC ligand-binding specificity between murine and human tetranectin
XX CC derivatives may be achieved. The present amino acid sequence represents a
XX CC human protein of the invention
XX SQ Sequence 186 AA;
Query Match 100.0%; Score 960; DB 5; Length 186;
Best Local Similarity 100.0%; Pred. No. 9.2e-93;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTDLAQEVALLEQQALQTVCLKGTVMKMC 60
DB 5 EPPTQPKKIVNAKDVVNTKMFELKSRDLTDLAQEVALLEQQALQTVCLKGTVMKMC 64
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 65 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 124
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 180
DB 125 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 184
QY 181 V 181
DB 185 V 185
RESULT 5
AAW94262
ID AAW94262 standard; protein; 197 AA.
XX
```



QY 1 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 60  
DB 22 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 81  
QY 61 FLAFTQTQTFHEASDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTQTFHEASDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENCAVLGSAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENCAVLGSAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

## RESULT 7

AAAM24034  
ID AAAM24034 standard; protein; 202 AA.  
AC AAAM24034;  
DT 12-OCT-2001 (first entry)  
XX Human EST encoded protein SEQ ID NO: 1559.  
DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;  
KW gene therapy; nutrition.

OS Homo sapiens.  
XX WO200154477-A2.  
PN 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US002687.  
XX 25-JAN-2000; 2000US-00491404.  
PR 17-JUL-2000; 2000US-00617746.  
PR 03-AUG-2000; 2000US-00631451.  
PR 15-SEP-2000; 2000US-00663870.  
XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Wehrman T;  
XX WPI; 2001-476164/51.  
DR N-PSDB; AAH98693.

PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use.

PS Claim 20; Page 1069; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention

XX Sequence 202 AA;

Query Match 100.0%; Score 960; DB 4; Length 202;  
Best Local Similarity 100.0%; Pred. No. 1e-92; Indels 0; Gaps 0;  
Matches 181; Conservative 0; Mismatches 0;

QY 1 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 60  
DB 22 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 81  
QY 61 FLAFTQTQTFHEASDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTQTFHEASDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENCAVLGSAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENCAVLGSAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

## RESULT 8

ADE09166  
ID ADE09166 standard; protein; 202 AA.  
AC ADE09166;  
DT 29-JAN-2004 (first entry)  
XX Novel protein-related contig polypeptide sequence #232.  
DE novel gene; novel protein; tissue marker; molecular weight marker;  
KW chromosome marker; genetic disorder; contig.  
XX Unidentified.  
OS WO2003054152-A2.  
XX 03-JUL-2003.  
XX 10-DEC-2002; 2002WO-US039555.  
XX 10-DEC-2001; 2001US-0339739P.  
PR 11-DEC-2001; 2001US-0339453P.  
PR 14-MAR-2002; 2002US-0365091P.  
PR 14-MAR-2002; 2002US-0365384P.  
PR 12-APR-2002; 2002US-0372381P.  
PR 12-APR-2002; 2002US-0372615P.  
PR 24-APR-2002; 2002US-00128558.  
PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX WPI; 2003-569235/53.

PT New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.

PS Disclosure; SEQ ID NO 2710; 1177pp; English.

XX The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence was used in the  
CC exemplification of the invention.

XX Sequence 202 AA;

Query Match 100.0%; Score 960; DB 7; Length 202;

Best Local Similarity 100.0%; Pred. No. 1e-92;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60  
DB 22 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 9  
ADE34332  
ID ADE34332 standard; protein; 202 AA.  
AC ADE34332;  
XX  
DT 29-JAN-2004 (first entry)  
DE Tetranectin #SEQ ID 13.  
XX Osteopathic; vaccine; osteoarthritis; gene modulation.  
KW Homo sapiens.  
OS  
XX WO2003048394-A1.  
PN 12-JUN-2003.  
XX  
PF 03-DEC-2002; 2002WO-US038407.  
XX  
PR 03-DEC-2001; 2001US-0337417P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA Carman J, Nadler SG, Bowen M, Neubauer M, Lu P;  
PI WPI; 2003-513771/48.  
DR N-PSDB; ADE34331.  
XX  
PT Identifying a compound that modulates the activity of osteoarthritis-associated gene or protein by determining whether the test compound modulates the activity of the gene or protein expressed in the cell contacted with the compound.  
XX  
PS Claim 10; Fig 11; 90pp; English.  
XX  
CC The invention relates to an assay for identifying a compound that modulates the activity of a gene or protein associated with osteoarthritis. The method of the invention comprises providing a cell expressing a gene or protein associated with osteoarthritis, contacting the cell with a test compound, and determining whether the test compound modulates the activity of the gene or protein. The method is useful for preparing a composition for treating osteoarthritis. Sequences given in ADE34362-ADE34355 represent osteoarthritis genes and proteins of the invention.  
XX  
SQ Sequence 202 AA;

Query Match  
Best Local Similarity 100.0%; Score 960; DB 7; Length 202;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60  
DB 22 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 10  
ADN04116  
ID ADN04116 standard; protein; 202 AA.  
AC ADN04116;  
XX  
DT 01-JUL-2004 (first entry)  
DE Antipsoriatic protein sequence #253.  
XX  
KW antipsoriatic; gene therapy; psoriasis; diagnosis.  
OS Homo sapiens.  
XX WO2004028479-A2.  
PN 08-APR-2004.  
XX  
PF 25-SEP-2003; 2003WO-US030907.  
XX  
PR 25-SEP-2002; 2002US-0414006P.  
XX (GETH ) GENENTECH INC.  
PA Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;  
XX  
DR WPI; 2004-305105/28.  
DR N-PSDB; ADN04115.  
XX  
PT New PRO nucleic acid or polypeptide, useful for preparing a pharmaceutical composition for diagnosing or treating psoriasis in a mammal.  
XX  
SQ Claim 9; SEQ ID NO 510; 3069pp; English.  
XX  
CC The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polypeptides of the invention.  
XX  
SQ Sequence 202 AA;

Query Match  
Best Local Similarity 100.0%; Score 960; DB 8; Length 202;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60  
DB 22 EPPTQPKKIVNAKDVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAOQPDGKGTENCAVLGAANGKWFDRKCRDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

DB 142 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 11  
AAW94261  
ID AAW94261 standard; protein; 228 AA.  
AC AAW94261;  
XX  
XX  
XX 20-MAR-2003 (revised)  
XX 26-APR-1999 (first entry)  
XX  
XX CIIH6FXTN123 fusion protein sequence.

KW Trimeric polypeptide; tetranectin trimerising structural element; TTSE;  
KW fusion protein; ligand binding structure; toxin; enzyme; cytokine;  
KW artificial antibody; pharmacokinetic; pharmacodynamic; gene therapy;  
KW transfection; imaging; tumour; human; tetranectin; CIIH6FXTN123.  
XX  
XX Homo sapiens.  
XX  
XX WO9856906-A1.  
XX  
XX 17-DEC-1998.  
XX  
XX 11-JUN-1998; 98WO-DK000245.  
XX  
XX 11-JUN-1997; 97DK-00000685.  
XX  
XX (THOG/) THOGERSEN H C.  
XX (ETZE/) ETZERODT M.  
XX (HOLT/) HOLTET T L.  
XX (GRAV/) GRAVERSEN N J H.  
XX (KAST/) KASTRUP J S.  
XX (NIEL/) NIELSEN B B.  
XX (LARS/) LARSEN I K.

XX Thogersen HC, Etzerodt M, Holtet TL, Graversen NJH, Kastrup JS;  
XX Nielsen BB, Larsen IK;  
XX WPI; 1999-080897/07.

XX New monomer polypeptide constructs for diagnosis and therapy - comprise a  
XX tetranectin trimerising structural element covalently linked to at least  
XX one heterologous moiety for providing functional activity.

XX Disclosure; Fig 6; 110pp; English.

XX The invention relates to the design of trimeric polypeptides using  
XX polypeptide structural elements derived from the tetranectin protein  
XX family. The trimeric polypeptides constructed as a monomer polypeptide  
XX construct comprise at least one tetranectin trimerising structural  
XX element (TTSE) which is covalently linked to at least one heterologous  
XX moiety, the TTSE being capable of forming a stable complex with 2 other  
XX TTSEs, with the proviso that the heterologous moiety is different from  
XX any of the fusion proteins CIIH6FXTN123, H6FXTN12, H6FXTN23  
XX (AAW94261 to AAW94264). The TTSE can be used for the construction of  
XX conjugates with heterologous moieties such as a ligand binding structure,  
XX a toxin, a detectable label, an in situ activatable substance, an enzyme,  
XX a radioactive moiety, a cytokine, a non-proteinaceous polymer, a photo  
XX cross-linking agent, or a group facilitating conjugation of the monomer  
XX polypeptide construct to a target. They can be used as vehicles for  
XX assembling antibody fragments into oligomeric or multivalent entities for  
XX generating chimeric artificial antibodies having preselected  
XX pharmacokinetic and/or pharmacodynamic properties. The constructs can be  
XX used for targeted gene therapy involving selective delivery of the  
XX material for transfection or infection of the specific population of  
XX cells. They can also be used for delivering a substance to a cell or

CC tissue or for delivering an imaging or toxin-conjugated antibody to a  
CC tumour. They can also be used for prevention or treating a disease or for  
CC diagnosis. The TTSE provides a stable structure which can act as a  
CC vehicle for a wide variety of conjugates. The present sequence represents  
CC a CIIH6FXTN123 fusion protein sequence. The specification claims that the  
CC heterologous moiety to which the TTSEs of the invention are linked to is  
CC specifically different from the present fusion protein sequence. (Updated  
CC on 20-MAR-2003 to correct PA field.)  
XX  
XX Sequence 228 AA;

Query Match 100.0%; Score 960; DB 2; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.2e-92;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLAGTKVHMKC 60  
DB 48 EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLAGTKVHMKC 107  
QY 61 FLAFTQTTFHEASEDCISRGGLSTPTQSGENDALYELVRSVGNENAIWGLNDMAAE 120  
DB 108 FLAFTQTTFHEASEDCISRGGLSTPTQSGENDALYELVRSVGNENAIWGLNDMAAE 167  
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 168 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 227  
QY 181 V 181  
DB 228 V 228

RESULT 12  
ABR48483  
ID ABR48483 standard; protein; 202 AA.  
XX  
XX AC ABR48483;  
XX  
XX DT 13-JUN-2003 (first entry)  
XX Human Plasminogen carrier protein, PLCP.  
XX  
XX KW Human; GENSET; therapeutic; therapy.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200294864-A2.  
XX  
XX PD 28-NOV-2002.  
XX  
XX PF 06-AUG-2001; 2001WO-IB001715.  
XX  
XX PR 25-MAY-2001; 2001US-0293574P.  
XX PR 15-JUN-2001; 2001US-0298698P.  
XX PR 29-JUN-2001; 2001US-030277P.  
XX PR 13-JUL-2001; 2001US-0305456P.  
XX  
XX PA (GEST ) GENSET.  
XX  
XX PI Bejanin S, Tanaka H;  
XX WPI; 2003-129412/12.  
XX  
XX DR N-PSDB; ACC51090.  
XX  
XX New GENSET polynucleotides and polypeptides, useful for preparing a  
XX composition for treating GENSET-related disorders and as reagents in  
XX assays to quantitatively determined levels of GENSET expression in  
XX biological samples.  
XX  
XX Claim 2; Page 454; 505pp; English.

XX The present invention relates to novel human GENSET coding sequences  
XX (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET  
CC

CC sequences are useful for preparing a composition for treating GENSET-  
 CC related disorders. They can also be used as markers for tissues in which  
 CC the corresponding protein is preferentially expressed, as molecular  
 CC weight markers on Southern gels, as chromosome markers or tags to  
 CC identify chromosomes, and as reagents in assays to quantitatively  
 CC determined levels of GENSET expression in biological samples  
 XX  
 SQ Sequence 202 AA;

Query Match 99.7%; Score 957; DB 6; Length 202;  
 Best Local Similarity 99.4%; Pred. No. 2.2e-92;  
 Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPPTQPKKIVNAKDDVNTKMFEEELKSRDLTLAQEVALLKEQQAQTVCVKGTKVHMKC 60  
 DB 22 EPPTQPKKIVNAKDDVNTKMFEEELKSRDLTLAQEVALLKEQQAQTVCVKGTKVHMKC 81  
 QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
 DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
 QY 121 GTWDMTGARIYKKNWETETITAPDGGKTENCALVSGAANGKWFDRKCRDQLPYICQFGI 180  
 DB 142 GTWDMTGARIYKKNWETETITAPDGGKTENCALVSGAANGKWFDRKCRDQLPYICQFGI 201  
 QY 181 V 181  
 DB 202 V 202

RESULT 13  
 ADD27211  
 ID ADD27211 standard; protein; 180 AA.  
 AC ADD27211;  
 DT 15-JAN-2004 (first entry)  
 DE Human adipocyte Selected Interacting Domain (SID) prey protein #675.  
 KW Adipocyte; protein-protein interaction; protein complex;  
 KW bait-prey complex; Selected Interacting Domain; SID; drug screening;  
 KW drug discovery; metabolic disease; obesity; lipodystrophy;  
 KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;  
 KW adipogenesis modulation; gene therapy; human.  
 OS Homo sapiens.  
 PN WO200290544-A2.  
 PD 14-NOV-2002.  
 PF 03-MAY-2002; 2002WO-EP006333.  
 PR 04-MAY-2001; 2001US-0288885P.  
 PA (HYBR-) HYBRIGENICS.  
 PA (LYNX-) LYNX THERAPEUTICS INC.  
 PI Legrain P, Whiteside S, Mao J, Khrebtukova I, Luo S;  
 DR WPI; 2003-111975/10.  
 DR N-PSDB; ADD27210.

New complex of bait and prey between two polypeptides or polynucleotides  
 encoding the two polypeptides of adipocytes, useful for selecting a  
 modulating compound that inhibits or activates protein-protein  
 interactions.

Claim 6; SEQ ID NO 675; 232pp; English.

The invention relates to a bait-prey complex between two adipocyte  
 polypeptides, or between two polynucleotides encoding adipocyte

polypeptides. The invention also relates to Selected Interacting Domain  
 (SID) polypeptides which interact with selected bait polypeptides;  
 polynucleotides encoding SID polypeptides; vectors comprising SID  
 polynucleotides; recombinant host cells comprising an adipocyte  
 compound which modulates interactions between adipocyte polypeptides;  
 compound which modulates interactions between adipocyte polypeptides;  
 adipocyte modulator compounds identified using the method; a  
 pharmaceutical composition comprising an adipocyte modulator, or a SID-  
 encoding vector or host cell; and a protein chip comprising adipocyte  
 bait polypeptides. The bait-prey complexes of the invention are useful  
 for selecting a compound that inhibits or activates protein-protein  
 interactions between adipocyte polypeptides. The modulatory compounds  
 identified can be used in the treatment of metabolic diseases such as  
 obesity, lipodystrophy and type 2 diabetes mellitus, and in the  
 modulation of adipogenesis. The present sequence represents a  
 specifically claimed adipocyte SID prey polypeptide of the invention.

SQ Sequence 180 AA;

Query Match 99.5%; Score 955; DB 7; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 3e-92;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PPTQPKKIVNAKDDVNTKMFEEELKSRDLTLAQEVALLKEQQAQTVCVKGTKVHMKCF 61  
 DB 1 PPTQPKKIVNAKDDVNTKMFEEELKSRDLTLAQEVALLKEQQAQTVCVKGTKVHMKCF 60  
 QY 62 LAFTQTKTFHEASEDCISRGGLTSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 121  
 DB 61 LAFTQTKTFHEASEDCISRGGLTSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
 QY 122 TWDMTGARIYKKNWETETITAPDGGKTENCALVSGAANGKWFDRKCRDQLPYICQFGI 181  
 DB 121 TWDMTGARIYKKNWETETITAPDGGKTENCALVSGAANGKWFDRKCRDQLPYICQFGI 180

RESULT 14  
 ABG80912

ID ABG80912 standard; protein; 186 AA.

AC ABG80912;

DT 29-NOV-2002 (first entry)

DE Human Phtlec protein.

KW Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;  
 KW beta-strand; connecting segment; 14loop region; tetranectin;  
 KW ligand-binding specificity; human.

OS Homo sapiens.

PN WO200248189-A2.

PD 20-JUN-2002.

PF 13-DEC-2001; 2001WO-DK000825.

PR 13-DEC-2000; 2000DK-00001872.

PR 28-FEB-2001; 2001US-027098P.

PA (BORE-) BOREAN PHARMA AS.

PI Etzerodt M, Holtet TL, Gravarsen NJH, Thogersen HC;

DR WPI; 2002-643278/69.

DR N-PSDB; ABS66594.

PT Protein comprising a variant of model C-type lectin-like domains (CTLD),  
 in which alpha helices, beta-strands, connecting segments are conserved  
 to maintain CTLD scaffold structure, while the loop region is altered.

PS Claim 14; Fig 14; 168pp; English.



XX The present invention relates to a new protein with scaffold structure of  
CC C-type lectin-like domains (CTLD). The invention comprises a variant of a  
CC model CTLD where alpha-helices and beta-strands and connecting segments  
CC are conserved such that scaffold structure of C-type lectin-like domains  
CC (CTLD) is substantially maintained, while the 14loop region is altered by  
CC amino acid substitution, deletion, insertion or their combination. The  
CC invention is useful for preparing a library of nucleotide sequences  
CC encoding related proteins by randomising part or all of the nucleic acid  
CC sequence encoding the loop region of its CTLD. The artificial CTLD  
CC protein products are preferable to antibody derivatives as each binding  
CC site is a single structurally autonomous protein domain. When used as  
CC components of compositions to be used for in vivo diagnostic or  
CC therapeutic purposes, artificial CTLD protein products constructed on the  
CC basis of human CTLDs are virtually identical to the corresponding natural  
CC CTLD protein already present in the body and are therefore less  
CC immunogenic to the patient. They also have a smaller size, and thus  
CC provide tissue penetration and distribution, as well as shorter half life  
CC in circulation. Since murine and human tetranectin are identical in  
CC structure, straightforward swapping of polypeptide segments defining  
CC ligand-binding specificity between murine and human tetranectin  
CC derivatives may be achieved. The present amino acid sequence represents a  
CC human protein of the invention  
XX  
SQ Sequence 186 AA;

Query Match 97.5%; Score 936; DB 5; Length 186;  
Best Local Similarity 98.3%; Pred. No. 3.2e-90;  
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKC 60  
DB 5 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKV 64  
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 65 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 124  
QY 121 GTWVDMTGARIAYKNWETEITAQPDGKGTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 125 GTWVDMTGTRIAYKNWETEITAQPDGKGTENCANVLSGAANGKWFDRCDQLPYICQFGI 184  
QY 181 V 181  
DB 185 V 185

RESULT 15  
ABG80908  
ID ABG80908 standard; protein; 188 AA.

XX AC ABG80908;

XX DT 29-NOV-2002 (first entry)

XX DE Human FX-htlec protein.

XX KW Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;  
KW beta-strand; connecting segment; 14loop region; tetranectin;  
KW ligand-binding specificity; human.

XX OS Homo sapiens.

XX PN WO200248189-A2.

XX PD 20-JUN-2002.

XX PF 13-DEC-2001; 2001WO-DK000825.

XX PR 13-DEC-2000; 2000DK-00001872.

XX PR 28-FEB-2001; 2001US-0272098P.

XX PA (BORE-) BOREAN PHARMA AS.

XX Etzerodt M, Holtet TL, Gravarsen NJH, Thogersen HC;  
XX WPI; 2002-643278/69.  
DR N-PSDB; ABS66587.  
XX Protein comprising a variant of model C-type lectin-like domains (CTLD),  
PT in which alpha helices, beta-strands, connecting segments are conserved  
PT to maintain CTLD scaffold structure, while the loop region is altered.  
XX  
PS Claim 46; Fig 6; 168pp; English.

XX The present invention relates to a new protein with scaffold structure of  
CC C-type lectin-like domains (CTLD). The invention comprises a variant of a  
CC model CTLD where alpha-helices and beta-strands and connecting segments  
CC are conserved such that scaffold structure of C-type lectin-like domains  
CC (CTLD) is substantially maintained, while the 14loop region is altered by  
CC amino acid substitution, deletion, insertion or their combination. The  
CC invention is useful for preparing a library of nucleotide sequences  
CC encoding related proteins by randomising part or all of the nucleic acid  
CC sequence encoding the loop region of its CTLD. The artificial CTLD  
CC protein products are preferable to antibody derivatives as each binding  
CC site is a single structurally autonomous protein domain. When used as  
CC components of compositions to be used for in vivo diagnostic or  
CC therapeutic purposes, artificial CTLD protein products constructed on the  
CC basis of human CTLDs are virtually identical to the corresponding natural  
CC CTLD protein already present in the body and are therefore less  
CC immunogenic to the patient. They also have a smaller size, and thus  
CC provide tissue penetration and distribution, as well as shorter half life  
CC in circulation. Since murine and human tetranectin are identical in  
CC structure, straightforward swapping of polypeptide segments defining  
CC ligand-binding specificity between murine and human tetranectin  
CC derivatives may be achieved. The present amino acid sequence represents a  
CC human protein of the invention  
XX  
SQ Sequence 188 AA;

Query Match 97.5%; Score 936; DB 5; Length 188;  
Best Local Similarity 98.3%; Pred. No. 3.2e-90;  
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKC 60  
DB 8 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKV 67  
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 68 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 127  
QY 121 GTWVDMTGARIAYKNWETEITAQPDGKGTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 128 GTWVDMTGTRIAYKNWETEITAQPDGKGTENCANVLSGAANGKWFDRCDQLPYICQFGI 187  
QY 181 V 181  
DB 188 V 188

Search completed: April 26, 2005, 11:57:37  
Job time : 167 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:54:54 ; Search time 140 Seconds  
(without alignments)  
430.251 Million cell updates/sec

Title: US-09-445-576A-7

Perfect score: 960

Sequence: 1 EPPTQPKKIVNAKKDVVNT.....KWFDRCRDQLPYICQFGIV 181

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	186	16	US-10-450-472-9
2	960	100.0	202	14	US-10-308-448-13
3	957	99.7	202	10	US-09-992-600A-62
4	957	99.7	202	10	US-09-924-340-62
5	957	99.7	202	10	US-09-992-095B-62
6	957	99.7	202	10	US-09-999-570-62
7	957	99.7	202	14	US-10-000-489-62
8	957	99.7	202	14	US-10-000-986-62
9	957	99.7	202	17	US-10-154-678-62
10	957	99.7	202	17	US-10-838-854-62
11	936	97.5	186	16	US-10-450-472-13
12	936	97.5	188	16	US-10-450-472-2
13	914	95.2	173	15	US-10-139-794-245
					Sequence 9, Appl
					Sequence 13, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 13, Appl
					Sequence 245, Appl

14	777	80.9	186	16	US-10-450-472-36	Sequence 36, Appl
15	777	80.9	187	16	US-10-450-472-29	Sequence 29, Appl
16	743	77.4	142	16	US-10-450-472-11	Sequence 11, Appl
17	723	75.3	143	16	US-10-450-472-4	Sequence 4, Appl
18	719	74.9	142	16	US-10-450-472-15	Sequence 15, Appl
19	618	64.4	143	16	US-10-450-472-31	Sequence 31, Appl
20	617	64.3	142	16	US-10-450-472-38	Sequence 38, Appl
21	448	46.7	89	16	US-10-450-472-57	Sequence 57, Appl
22	422.5	44.0	197	9	US-09-745-288-99	Sequence 99, Appl
23	422.5	44.0	197	14	US-10-177-293-55	Sequence 55, Appl
24	422.5	44.0	197	15	US-10-453-919-99	Sequence 99, Appl
25	422.5	44.0	206	9	US-09-989-723-403	Sequence 403, Appl
26	422.5	44.0	206	9	US-09-989-723-403	Sequence 403, Appl
27	422.5	44.0	206	9	US-09-989-279-403	Sequence 403, Appl
28	422.5	44.0	206	9	US-09-989-727-403	Sequence 403, Appl
29	422.5	44.0	206	9	US-09-989-731-403	Sequence 403, Appl
30	422.5	44.0	206	9	US-09-989-732-403	Sequence 403, Appl
31	422.5	44.0	206	9	US-09-991-073-403	Sequence 403, Appl
32	422.5	44.0	206	9	US-09-990-442-403	Sequence 403, Appl
33	422.5	44.0	206	9	US-09-991-163-403	Sequence 403, Appl
34	422.5	44.0	206	9	US-09-989-293A-403	Sequence 403, Appl
35	422.5	44.0	206	9	US-09-993-604-403	Sequence 403, Appl
36	422.5	44.0	206	9	US-09-990-456-403	Sequence 403, Appl
37	422.5	44.0	206	9	US-09-989-721-403	Sequence 403, Appl
38	422.5	44.0	206	9	US-09-992-598-403	Sequence 403, Appl
39	422.5	44.0	206	9	US-09-938-418-10	Sequence 10, Appl
40	422.5	44.0	206	9	US-09-989-735-403	Sequence 403, Appl
41	422.5	44.0	206	9	US-09-990-444-403	Sequence 403, Appl
42	422.5	44.0	206	9	US-09-991-181-403	Sequence 403, Appl
43	422.5	44.0	206	9	US-09-989-730-403	Sequence 403, Appl
44	422.5	44.0	206	9	US-09-990-436-403	Sequence 403, Appl
45	422.5	44.0	206	9	US-09-993-687-403	Sequence 403, Appl

ALIGNMENTS

RESULT 1

US-10-450-472-9  
; Sequence 9, Application US/10450472  
; Publication No. US20040132094A1  
; GENERAL INFORMATION:  
; APPLICANT: Borean Pharma A/S  
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure  
; FILE OF INVENTION: of C-type lectin-like domains  
; FILE REFERENCE: BOR0003/WO  
; CURRENT APPLICATION NUMBER: US/10/450,472  
; CURRENT FILING DATE: 2003-12-08  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9  
; LENGTH: 186  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-450-472-9

Query Match	100.0%	Score	960;	DB	16;	Length	186;
Best Local Similarity	100.0%;	Pred. No.	4.3e-92;				
Matches	181;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTVHKMC	60				
Db	5	EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTVHKMC	64				
Qy	61	FLAFTQTTTFHEASDCISRGSTLSTPTQSGENDALYELVRSVGNRAEILWGLNDMAAE	120				
Db	65	FLAFTQTTTFHEASDCISRGSTLSTPTQSGENDALYELVRSVGNRAEILWGLNDMAAE	124				
Qy	121	GTWVDMTCARLAYKNWETEITTAQPDGGKTENCANVLSGAANGKWFDRCDRLPYICQFGI	180				
Db	125	GTWVDMTCARLAYKNWETEITTAQPDGGKTENCANVLSGAANGKWFDRCDRLPYICQFGI	184				
Qy	181	V	181				

```
Db      185 V 185
|
RESULT 2
US-10-308-448-13
; Sequence 13, Application US/10308448
; Publication No. US20030170743A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE DEVELOPMENT
; FILE REFERENCE: D0189
; CURRENT APPLICATION NUMBER: US/10/308,448
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/337,417
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-308-448-13

Query Match      100.0%; Score 960; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.8e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
Db      22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY      61 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db      82 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY      121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
Db      142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY      181 V 181
|
Db      202 V 202
|
RESULT 3
US-09-992-600A-62
; Sequence 62, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-992-600A-62

Query Match      99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
Db      22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY      61 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db      82 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY      121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
Db      142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY      181 V 181
|
Db      202 V 202
|
RESULT 4
US-09-924-340-62
; Sequence 62, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-924-340-62

Query Match      99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
Db      22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY      61 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db      82 FLAFTQTKTFHASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY      121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
Db      142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY      181 V 181
|
Db      202 V 202
|
```

```
QY 181 V 181
Db 202 V 202

RESULT 5
US-09-992-095B-62
; Sequence 62, Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-09-999-570-62

Query Match 99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9,9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 60
Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 81
QY 61 FLAFTQTKTFHEASDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 FLAFTQTKTFHEASDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 180
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
Db 202 V 202

RESULT 6
US-09-999-570-62
; Sequence 62, Application US/09999570
; Publication No. US20030170628A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: G-091US08DIV
; CURRENT APPLICATION NUMBER: US/09/999,570
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/924,340

Query Match 99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9,9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 60
Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 81
QY 61 FLAFTQTKTFHEASDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 FLAFTQTKTFHEASDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 180
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
Db 202 V 202

RESULT 7
US-10-000-489-62
; Sequence 62, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-489-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
   |||||
Db 22 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
   |||||
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
   |||||
Db 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
   |||||
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
   |||||
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
   |||||
QY 181 V 181
   |
Db 202 V 202

RESULT 8
US-10-000-986-62
; Sequence 62, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US 10/000,986
; PRIOR FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-986-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
   |||||
Db 22 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
   |||||
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
   |||||
Db 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
   |||||
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
   |||||
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
   |||||
QY 181 V 181
   |
Db 202 V 202

RESULT 9
US-10-154-678-62
; Sequence 62, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US 10/154,678
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -21..-1
US-10-154-678-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
   |||||
Db 22 EPPTQPKKIVNAKDVVNTKMFEEELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
   |||||
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
   |||||
Db 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
   |||||
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
   |||||
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
   |||||
QY 181 V 181
   |
Db 202 V 202

RESULT 10
US-10-838-854-62
; Sequence 62, Application US/10838854
; Publication No. US20050026182A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US 10/838,854
; PRIOR FILING DATE: 2004-05-03
; PRIOR APPLICATION NUMBER: US 10/000,489
; PRIOR FILING DATE: 2001-11-14
```

PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: Jpatent  
SEQ ID NO 62  
LENGTH: 202  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: 1..21  
US-10-838-854-62

Query Match 99.7%; Score 957; DB 17; Length 202;  
Best Local Similarity 99.4%; Pred. No. 9.9e-92;  
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 60  
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 11  
US-10-450-472-13  
Sequence 13, Application US/10450472  
Publication No. US20040132094A1  
GENERAL INFORMATION:  
APPLICANT: Boreau Pharma A/S  
TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure  
FILE REFERENCE: BOR0003/WO  
CURRENT APPLICATION NUMBER: US/10/450,472  
CURRENT FILING DATE: 2003-12-08  
NUMBER OF SEQ ID NOS: 91  
SEQ ID NO 13  
LENGTH: 186  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-450-472-13

Query Match 97.5%; Score 936; DB 16; Length 186;  
Best Local Similarity 98.3%; Pred. No. 1.4e-89;  
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 60  
DB 5 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 64  
QY 61 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120

DB 65 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 124  
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 125 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 184  
QY 181 V 181  
DB 185 V 185

RESULT 12  
US-10-450-472-2  
Sequence 2, Application US/10450472  
Publication No. US20040132094A1  
GENERAL INFORMATION:  
APPLICANT: Boreau Pharma A/S  
TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure  
FILE REFERENCE: BOR0003/WO  
CURRENT APPLICATION NUMBER: US/10/450,472  
CURRENT FILING DATE: 2003-12-08  
NUMBER OF SEQ ID NOS: 91  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2  
LENGTH: 188  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-450-472-2

Query Match 97.5%; Score 936; DB 16; Length 188;  
Best Local Similarity 98.3%; Pred. No. 1.4e-89;  
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 60  
DB 8 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVCLKGTVMHMC 67  
QY 61 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
DB 68 FLAFTQTKTFHEASEDCISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 127  
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 128 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 187  
QY 181 V 181  
DB 188 V 188

RESULT 13  
US-10-139-794-245  
Sequence 245, Application US/10139794  
Publication No. US20030232421A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS, LYNX THERAPEUTICS INC.  
APPLICANT: Pierre Legrain, Simon Whiteside, Jen-I Mao, Irina Khrebukova, Shujun Luo  
TITLE OF INVENTION: Protein-Protein Interactions In Adipocyte Cells (3)  
FILE REFERENCE: B4883A  
CURRENT APPLICATION NUMBER: US/10/139,794  
CURRENT FILING DATE: 2002-05-06  
PRIOR APPLICATION NUMBER: US 60/288,885  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 2930  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 245  
LENGTH: 173  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
OTHER INFORMATION: Translation of SEQ ID NO:242  
US-10-139-794-245

```
Query Match      95.2%; Score 914; DB 15; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.5e-87;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 KIVNAKDVNTKMFELKSRDLTLAQEVALLKEQALQTVCLKGPKVHKMCFLLAFTQTK 68
DB 1 KIVNAKDVNTKMFELKSRDLTLAQEVALLKEQALQTVCLKGPKVHKMCFLLAFTQTK 60

QY 69 TTHEASEDCISRGGLTSTPQTSSENDALYEYLRSQVGNAAEIVLGLNDMAAEGTWVDMTG 128
DB 61 TTHEASEDCISRGGLTSTPQTSSENDALYEYLRSQVGNAAEIVLGLNDMAAEGTWVDMTG 120

QY 129 ARIAYKNWETEITAOQPDGKTCENCAVLGAANGKWFDRKRCRDLQPYICQFQIV 181
DB 121 ARIAYKNWETEITAOQPDGKTCENCAVLGAANGKWFDRKRCRDLQPYICQFQIV 173

RESULT 14
US-10-450-472-36
; Sequence 36, Application US/10450472
; Publication No. US20040132094A1
; GENERAL INFORMATION:
; APPLICANT: Borean Pharma A/S
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
; FILE REFERENCE: BOR0003/WO
; CURRENT APPLICATION NUMBER: US/10/450,472
; CURRENT FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-450-472-36

Query Match      80.9%; Score 777; DB 16; Length 186;
Best Local Similarity 80.1%; Pred. No. 6.1e-73;
Matches 145; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQALQTVCLKGTVHKMC 60
DB 5 ESPTPKAKAANAADLVSSKMFELKSRDLTLAQEVALLKEQALQTVCLKGTVHKMC 64

QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTSSENDALYEYLRSQVGNAAEIVLGLNDMAAE 120
DB 65 LLAFTQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHSVGNDAEIVLGLNDMAAE 124

QY 121 GTWVDMTGARIAYKNWETEITAOQPDGKTCENCAVLGAANGKWFDRKRCRDLQPYICQFQI 180
DB 125 GAWVDMTGTLAYKNWETEITTPDGGKAENCAALSGAANGKWFDRKRCRDLQPYICQFAI 184

QY 181 V 181
DB 185 V 185

RESULT 15
US-10-450-472-29
; Sequence 29, Application US/10450472
; Publication No. US20040132094A1
; GENERAL INFORMATION:
; APPLICANT: Borean Pharma A/S
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
; FILE REFERENCE: BOR0003/WO
; CURRENT APPLICATION NUMBER: US/10/450,472
; CURRENT FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 187
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-450-472-29

Query Match      80.9%; Score 777; DB 16; Length 187;
Best Local Similarity 80.1%; Pred. No. 6.1e-73;
Matches 145; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQALQTVCLKGTVHKMC 60
DB 7 ESPTPKAKAANAADLVSSKMFELKSRDLTLAQEVALLKEQALQTVCLKGTVHKMC 66

QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTSSENDALYEYLRSQVGNAAEIVLGLNDMAAE 120
DB 67 LLAFTQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHSVGNDAEIVLGLNDMAAE 126

QY 121 GTWVDMTGARIAYKNWETEITAOQPDGKTCENCAVLGAANGKWFDRKRCRDLQPYICQFQI 180
DB 127 GAWVDMTGTLAYKNWETEITTPDGGKAENCAALSGAANGKWFDRKRCRDLQPYICQFAI 186

QY 181 V 181
DB 187 V 187

Search completed: April 26, 2005, 12:12:13
Job time : 142 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 42 Seconds  
(without alignments)  
321.702 Million cell updates/sec

Title: US-09-445-576A-7  
Perfect score: 960  
Sequence: 1 BPPTQPKKIIVNAKDVNT.....KWFDKRCRDQLPYICQFQIV 181

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	960	100.0	202	1	US-08-469-486-56
2	960	100.0	202	2	US-08-469-658-56
3	957	99.7	202	4	US-10-000-489-62
4	422.5	44.0	197	4	US-09-602-877A-99
5	222.5	23.2	328	4	US-09-065-040-12
6	221.5	23.1	328	4	US-09-065-040-8
7	220.5	23.0	323	4	US-09-065-040-4
8	220.5	23.0	323	4	US-09-949-016-6294
9	189	19.7	59	4	US-09-513-999C-4342
10	186.5	19.4	1257	1	US-08-340-428B-49
11	185.5	19.3	248	3	US-09-198-603C-2
12	185.5	19.3	248	4	US-09-600-932-27
13	185.5	19.3	248	4	US-09-949-016-5907
14	185.5	19.3	263	4	US-09-949-016-9370
15	178.5	18.6	161	3	US-09-011-735-6
16	178.5	18.6	351	3	US-09-011-735-1
17	178.5	18.6	351	3	US-09-029-156-1
18	177	18.4	912	5	PCT-US95-03747-2
19	172.5	18.0	375	4	US-09-600-932-29
20	166	17.3	287	1	US-08-365-103B-4
21	166	17.3	300	1	US-08-365-103B-6
22	166	17.3	327	1	US-08-365-103B-2
23	165	17.2	1479	3	US-08-840-062-4
24	163	17.0	1456	4	US-09-976-594-168
25	160	16.7	1479	3	US-08-840-062-2
26	152	15.8	174	2	US-08-401-530A-5
27	152	15.8	174	2	US-08-709-662-5

28	151	15.7	1487	3	US-08-840-062-7	Sequence 7, Appli
29	149.5	15.6	1326	4	US-09-949-016-6806	Sequence 6806, Ap
30	148.5	15.5	1339	4	US-09-949-016-10448	Sequence 10448, A
31	148.5	15.5	1478	4	US-09-949-016-8315	Sequence 8315, Ap
32	148	15.4	125	3	US-08-722-136A-7	Sequence 7, Appli
33	148	15.4	125	5	PCT-US95-04258-7	Sequence 7, Appli
34	148	15.4	166	2	US-08-729-103-4	Sequence 4, Appli
35	148	15.4	174	4	US-09-949-016-10686	Sequence 10686, A
36	148	15.4	2409	6	5180808-2	Patent No. 5180808
37	148	15.4	2409	6	5180808-2	Patent No. 5180808
38	142	14.8	108	6	5514582-16	Patent No. 5514582
39	142	14.8	108	6	5514582-16	Patent No. 5514582
40	141	14.7	455	4	US-09-866-028-50	Sequence 50, Appl
41	141	14.7	455	4	US-09-944-457-50	Sequence 50, Appl
42	140.5	14.6	187	4	US-09-535-521-17	Sequence 17, Appl
43	140.5	14.6	208	4	US-09-535-521-20	Sequence 20, Appl
44	140.5	14.6	292	4	US-09-535-521-2	Sequence 2, Appli
45	140.5	14.6	292	4	US-09-535-521-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-469-486-56  
; Sequence 56, Application US/08469486  
; Patent No. 5739281  
; GENERAL INFORMATION:  
; APPLICANT: Thøgersen, Hans Christian  
; APPLICANT: Holtet, Thor Las  
; APPLICANT: Etzerodt, Michael  
; TITLE OF INVENTION: Improved method for the refolding of  
; TITLE OF INVENTION: proteins  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version  
SOFTWARE: #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,486  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/192,060  
FILING DATE: February 4, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06363/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 542 5070  
TELEFAX: 617 542 8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
TYPE: amino acid  
LENGTH: 202 amino acids  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-486-56

Query Match 100.0%; Score 960; DB 1; Length 202;  
Best Local Similarity 100.0%; Pred. No. 5.3e-102;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 60  
DB 22 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 2  
US-08-469-658-56  
; Sequence 56, Application US/08469658  
; Patent No. 5917018  
; GENERAL INFORMATION:  
; APPLICANT: Th egersen, Hans Christian  
; APPLICANT: Holtet, Thor Las  
; APPLICANT: Etzerodt, Michael  
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version  
; SOFTWARE: #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,658  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/192,060  
; FILING DATE: February 4, 1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06363/002002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 542 5070  
; TELEFAX: 617 542 8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 202 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-469-658-56

Query Match 100.0%; Score 960; DB 2; Length 202;  
Best Local Similarity 100.0%; Pred. No. 5.3e-102;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 60

DB 22 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 3  
US-10-000-489-62  
; Sequence 62, Application US/10000489  
; Patent No. 6794363  
; GENERAL INFORMATION:  
; APPLICANT: Benjanin, Stephanie  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 62  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: 1..21  
US-10-000-489-62

Query Match 99.7%; Score 957; DB 4; Length 202;  
Best Local Similarity 99.4%; Pred. No. 1.2e-101;  
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 60  
DB 22 EPPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEKQALQTVCLKGTKVHMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRSQVNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 180  
DB 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENC AVLSGAANGKWFDRCDQLPYICQFGI 201  
QY 181 V 181  
DB 202 V 202

RESULT 4  
US-09-602-877A-99

; Sequence 99, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602,877A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 99  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-602-877A-99

Query Match 44.0%; Score 422.5; DB 4; Length 197;  
Best Local Similarity 51.3%; Pred. No. 2.9e-40;  
Matches 80; Conservative 27; Mismatches 46; Indels 3; Gaps 1;  
  
QY 25 ELKSLDITLAORVALLKEQALQTVCLKGTQVHKMCFLAFTQTKTFHEASEDCISRGCTL 84  
DB 43 DLKTOIEKLWTEVNALKEIQALQTVCLKGTQVHKMCFLAFTQTKTFHEASEDCISRGCTL 102  
  
QY 85 SPOTGSENDALEYLRQSVGNEABITWGLNDMAAEGTWDMTGARIAYKNWETEITAPQ 144  
DB 103 VIPRNSDEINADQYKSLPGVNDVFWLGINDMVTEGKFDVNGIAISFLNWD---RAQP 159  
  
QY 145 DGGKTENCALVGAANGKWFDRCDLPYICQFI 180  
DB 160 NGGKRECVLFSQAGKQWSDACRSKRYICEFTI 195

## RESULT 5

US-09-065-040-12  
; Sequence 12, Application US/09065040  
; Patent No. 6541217  
; GENERAL INFORMATION:  
; APPLICANT: Hiraoka, Atsunobu  
; APPLICANT: Sugimura, Atsushi  
; APPLICANT: Mio, Hiroyuki  
; TITLE OF INVENTION: HEMATOPOIETIC STEM CELL GROWTH FACTOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
; ADDRESSEE: DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/065,040  
; FILING DATE: 27-APR-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 262252/1996  
; FILING DATE: 27-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 087242/1997  
; FILING DATE: 24-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP97/02349  
; FILING DATE: 07-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fordis, Jean B.

; NAME: Fordis, Jean B.  
; REGISTRATION NUMBER: 32,984  
; REFERENCE/DOCKET NUMBER: 04853.0026-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-065-040-12

Query Match 23.2%; Score 222.5; DB 4; Length 328;  
Best Local Similarity 28.6%; Pred. No. 6.1e-17;  
Matches 55; Conservative 33; Mismatches 73; Indels 31; Gaps 5;  
  
QY 17 VVNTKMF-----ELKSLDITLAORVALLKEQALQTV-----CLKGTQVHKMCFLA 63  
DB 136 VLDTRVVELTQGLRRLRDAASDTRDSVQALKEVQVRSEQHGRLEGCLKGLRLGHKCFLL 195  
  
QY 64 FTQTKTFHEASEDCISRGCTLSTPQTGSENDALEYLRQSVG-NEABITWGLNDMAAEGT 122  
DB 196 SRDFETQAAQAQRCVARGGSLAQPADRQOMDALSRYLRAALAPYWPVWLVGVDHRSRSEGL 255  
  
QY 123 WVDMTGARIAYKNWETEITATA-----QPDGKKTENCALVGAANGKWFDRK 166  
DB 256 YLFENGQVRVSPFAWHRALSPESGAHPLSPDQPNGGILENC-VAQASDDGSGWWDH 314  
  
QY 167 RCRDLPYICQF 178  
DB 315 DCERRLYFVCFE 326

## RESULT 6

US-09-065-040-8  
; Sequence 8, Application US/09065040  
; Patent No. 6541217  
; GENERAL INFORMATION:  
; APPLICANT: Hiraoka, Atsunobu  
; APPLICANT: Sugimura, Atsushi  
; APPLICANT: Mio, Hiroyuki  
; TITLE OF INVENTION: HEMATOPOIETIC STEM CELL GROWTH FACTOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
; ADDRESSEE: DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/065,040  
; FILING DATE: 27-APR-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 262252/1996  
; FILING DATE: 27-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 087242/1997  
; FILING DATE: 24-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP97/02349  
; FILING DATE: 07-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fordis, Jean B.

REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04853.0026-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 328 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-065-040-8

Query Match 23.1%; Score 221.5; DB 4; Length 328;  
Best Local Similarity 27.8%; Pred. No. 7.9e-17;  
Matches 55; Conservative 36; Mismatches 76; Indels 31; Gaps 5;

QY 11 VNAKDVNTKPE-----ELKRLDTLAQEVALLKEQO-----ALQTVCLGKTKVH 57  
Db 130 LHVRLHLVDTVVVLTQGLRQLRDAASDTRDSVQALKEVDRAEQSHGRLEGCLKGLRLG 189

QY 58 MKCFLAFTQTKTPEASEDCISRGGLTSTPQTGSENDALYEYLRQSVG-NEAEIWLGLND 116  
Db 190 HKCFLLSRFETQAAQARCKARGGSLAQPADRQOMDALSRYLRAALAPYNNPVLGVHD 249

QY 117 MAAGTVDVMTGARIAYKNWETEITA-----QPDGKTKENCAVLSGAAN 160  
Db 250 RRSGLYLFENGQVRSFFFAWHRSFLESQAQPSAATHPLSPDPNGVGLENC-VAQASDD 308

QY 161 GKWFDRCDLPYICQF 178  
Db 309 GSWWDHDCERRLYVCEF 326

RESULT 7  
US-09-065-040-4  
Sequence 4, Application US/09065040  
Patent No. 6541217  
GENERAL INFORMATION:  
APPLICANT: Hiraoka, Atsunobu  
APPLICANT: Sugimura, Atsushi  
APPLICANT: Mio, Hiroyuki  
TITLE OF INVENTION: HEMATOPOIETIC STEM CELL GROWTH FACTOR  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
ADDRESSEE: DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/065,040  
FILING DATE: 27-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 262252/1996  
FILING DATE: 27-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 087242/1997  
FILING DATE: 24-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP97/02349  
FILING DATE: 07-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984

REFERENCE/DOCKET NUMBER: 04853.0026-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 323 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-065-040-4

Query Match 23.0%; Score 220.5; DB 4; Length 323;  
Best Local Similarity 29.5%; Pred. No. 1e-16;  
Matches 57; Conservative 37; Mismatches 70; Indels 29; Gaps 7;

QY 4 TOKPKKIVNAKKDVVN-TKMFELKSRDLDTLAQEVALLKEQOALQTVCLKGTKVHMKCFL 62  
Db 140 TQGLRQLRNAAGDTRDAVQALQEAQGRAB-----REHGRLEG-CLKGLRLGHKCF 189

QY 63 AFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVG-NEAEIWLGLNDMAE 121  
Db 190 LSRDFEAQAAQARCTARGGSLAQPADRQOMDALSRYLRAALAPYNNPVLGVHDRAE 249

QY 122 TWVDMTGARIAYKNW-----ETEITA-----OPDGKTKENCAVLSGAANGK 165  
Db 250 LYLFPNGQVRSFFFAWHRSRPELGAQPSASPHPLSPDPNGVGLENC-VAQASDDG 308

QY 166 KCRDLPYICQF 178  
Db 309 HDCQRRLYVCEF 321

RESULT 8  
US-09-949-016-6294  
Sequence 6294, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6294  
LENGTH: 323  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-6294

Query Match 23.0%; Score 220.5; DB 4; Length 323;  
Best Local Similarity 29.5%; Pred. No. 1e-16;  
Matches 57; Conservative 37; Mismatches 70; Indels 29; Gaps 7;

QY 4 TOKPKKIVNAKKDVVN-TKMFELKSRDLDTLAQEVALLKEQOALQTVCLKGTKVHMKCFL 62  
Db 140 TQGLRQLRNAAGDTRDAVQALQEAQGRAB-----REHGRLEG-CLKGLRLGHKCF 189

QY 63 AFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVG-NEAEIWLGLNDMAE 121  
Db 190 LSRDFEAQAAQARCTARGGSLAQPADRQOMDALSRYLRAALAPYNNPVLGVHDRAE 249

QY 122 TWVDMTGARIAYKNW-----ETEITA-----OPDGKTKENCAVLSGAANGK 165  
Db 250 LYLFPNGQVRSFFFAWHRSRPELGAQPSASPHPLSPDPNGVGLENC-VAQASDDG 308

```
QY 166 KRCRDLPYICOF 178
Db 309 HDCORRLYYVCF 321

RESULT 9
US-09-513-999C-4342
; Sequence 4342, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4342
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: -21...-1
; OTHER INFORMATION: score 11.2
; OTHER INFORMATION: seq LILCLFSLLTQVTT/EP
US-09-513-999C-4342

Query Match 19.7%; Score 189; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.1e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 38; Conservative 0;

QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVA 38
Db 22 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVA 59

RESULT 10
US-08-340-428B-49
; Sequence 49, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broadway and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,428B
; FILING DATE: 14 No. 5648465ember 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/922,911
; FILING DATE: 03 August 1992

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: Margolis-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-428B-49

Query Match 19.4%; Score 186.5; DB 1; Length 1257;
Best Local Similarity 32.7%; Pred. No. 6.5e-12; Mismatches 20; Indels 27; Gaps 8;
Matches 48; Conservative 20;

QY 49 VCLKGT-----KVMKCFLAFTQTTFHEASEDCISRGCTLTPTQSGENDALYEYL 100
Db 1021 LCEKDETCGDRGWHKFGQHCYRYFAHRAWEDAERDCRRAGHLTSVHSPEHKFI---- 1076
QY 101 RQSVGNRAEIMGLNDMAAG--TWVDMTGARIAYKNWETEITAQPD-----GGKTENCAY 154
Db 1077 -NSFGHE-NSWIGLNDRTVERDFQWTDNTG--LQYENWREK---QPDNFFAGG--EDCVV 1127
QY 155 LSGAANGKWFDRCDQLPYICQFIV 181
Db 1128 MVAHENGWRNDVPCNYNLPYVCKKGTV 1154

RESULT 11
US-09-198-603C-2
; Sequence 2, Application US/09198603C
; Patent No. 6337193
; GENERAL INFORMATION:
; APPLICANT: TULLY, Raymond E.
; APPLICANT: CALTAGIRONE, G. Thomas
; APPLICANT: MOYER, Shawn S.
; APPLICANT: RONNING, Michael T.
; TITLE OF INVENTION: EXPRESSION OF MANNOSE-BINDING PROTEIN IN METHYLOTROPHIC
; TITLE OF INVENTION: YEAST
; FILE REFERENCE: A7290
; CURRENT APPLICATION NUMBER: US/09/198,603C
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-198-603C-2

Query Match 19.3%; Score 185.5; DB 3; Length 248;
Best Local Similarity 30.2%; Pred. No. 7.1e-13; Mismatches 28; Indels 21; Gaps 5;
Matches 45; Conservative 28;

QY 42 EQQALQTVCLK-----GTYVMKCFLAFTQTKTFHEASEDCISRGCTLTPTQSGE 92
Db 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOASVATPRNAE 170
QY 93 NDALYEYLRQSVGNRAEIMGLNDMAAGTWVDMTGARIAYKNWETEITAQPD-DGSKTEN 151
Db 171 NGAIONLIKE-----EAPLGITDEKTEGQFVDLTGNRLTYTNWN---EGEPNAGSDED 221
QY 152 CAVLSGAANGKWFDRCDQLPYICQFIV 180
Db 222 CVLL--LKNQWQNDVPCSTSHLAVCEPPI 248
```

## RESULT 12

US-09-600-932-27  
; Sequence 27, Application US/09600932  
; Patent No. 6787639  
; GENERAL INFORMATION:  
; APPLICANT: Wakamiya, No. 6787639utaka  
; TITLE OF INVENTION: NOVEL COLLECTIN  
; FILE REFERENCE: 19036/36615  
; CURRENT APPLICATION NUMBER: US/09/600,932  
; CURRENT FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: PCT/JP98/03328  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 27  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: mannan-binding protein (MBP)  
US-09-600-932-27

Query Match 19.3%; Score 185.5; DB 4; Length 248;  
Best Local Similarity 30.2%; Pred. No. 7.1e-13;  
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;  
QY 42 EQALQTVCLK-----GTVHMKCFLAFTQTKTFHEASEDCISRGGLSTPQTGSE 92  
DB 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 170  
QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETETIAQP-DGGKTEN 151  
DB 171 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 221  
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180  
DB 222 CVLL--LKNQWMDVPCSTSHLAVCEPFI 248

## RESULT 13

US-09-949-016-5907  
; Sequence 5907, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5907  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-5907

Query Match 19.3%; Score 185.5; DB 4; Length 248;  
Best Local Similarity 30.2%; Pred. No. 7.1e-13;  
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;  
QY 42 EQALQTVCLK-----GTVHMKCFLAFTQTKTFHEASEDCISRGGLSTPQTGSE 92  
DB 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 170

QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETETIAQP-DGGKTEN 151  
DB 171 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 221  
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180  
DB 222 CVLL--LKNQWMDVPCSTSHLAVCEPFI 248

## RESULT 14

US-09-949-016-9370  
; Sequence 9370, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9370  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9370

Query Match 19.3%; Score 185.5; DB 4; Length 263;  
Best Local Similarity 30.2%; Pred. No. 7.8e-13;  
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;  
QY 42 EQALQTVCLK-----GTVHMKCFLAFTQTKTFHEASEDCISRGGLSTPQTGSE 92  
DB 126 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 185  
QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETETIAQP-DGGKTEN 151  
DB 186 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 236  
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180  
DB 237 CVLL--LKNQWMDVPCSTSHLAVCEPFI 263

## RESULT 15

US-09-011-735-6  
; Sequence 6, Application US/09011735B  
; Patent No. 6110708  
; GENERAL INFORMATION:  
; APPLICANT: Wakamiya, No. 6110708utaka  
; TITLE OF INVENTION: Recombinant Conglutinin and Producing Method Thereof  
; FILE REFERENCE: 19036/34548  
; CURRENT APPLICATION NUMBER: US/09/011,735B  
; CURRENT FILING DATE: 1998-05-22  
; EARLIER APPLICATION NUMBER: JP 7-209698  
; EARLIER FILING DATE: 1995-08-17  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 161  
; TYPE: PRT  
; ORGANISM: Bovine  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (2)

Search completed: April 26, 2005, 11:54:46  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:11 ; Search time 176 Seconds  
(without alignments)  
526.627 Million cell updates/sec

Title: US-09-445-576A-7  
Perfect score: 960  
Sequence: 1 EPPTQPKKIVNAKDVNT.....KWFDRCRQLPYICQFGIV 181

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	960	100.0	202	1	TETN_HUMAN	P05452 homo sapien
2	792	82.5	202	1	TETN_MOUSE	P43025 mus musculus
3	792	82.5	202	2	Q8CFZ6	Q8cfz6 mus musculus
4	725	75.5	133	2	Q68DS3	Q68ds3 homo sapien
5	720.5	75.1	201	2	Q9DDA4	Q9dda4 gallus gall
6	455	47.4	193	2	Q68KU1	Q68kul xenopus lae
7	431.5	44.9	197	2	Q28008	Q28008 bos taurus
8	430.5	44.8	196	2	Q9EPW4	Q9epw4 mus musculus
9	422.5	44.0	197	1	CLB1_HUMAN	O75596 homo sapien
10	422.5	44.0	206	2	Q6UXF5	Q6uxf5 homo sapien
11	408	42.5	166	1	TETN_CARSP	P26258 carcharhinu
12	222.5	23.2	328	1	SCGF_RAT	O88201 rattus norv
13	221.5	23.1	328	1	SCGF_MOUSE	O88200 mus musculus
14	220.5	23.0	323	1	SCGF_HUMAN	Q9Y240 homo sapien
15	207.5	21.6	369	1	PSPD_BOVIN	P35246 bos taurus
16	199.5	20.8	369	2	O863A1	Q863a1 bos taurus
17	188.5	19.6	652	2	Q8WSX1	O8wsx1 dugesia tig
18	187.5	19.5	374	1	PSPD_MOUSE	P50404 mus musculus
19	186.5	19.4	244	1	MABC_MOUSE	P41317 mus musculus
20	186.5	19.4	1257	1	PGCN_RAT	P55067 rattus norv
21	185.5	19.3	248	1	MABC_HUMAN	P11226 homo sapien
22	185.5	19.3	248	1	MABC_HYLCO	Q66588 hylobates c
23	184.5	19.2	248	1	MABC_GORGO	Q66560 gorilla gor
24	183.5	19.1	248	1	MABC_PANTR	Q66563 pan troglod
25	183.5	19.1	1290	2	Q9W6E1	Q9w6e1 gallus gall
26	183	19.1	246	2	Q919Q7	Q919q7 carassius a
27	182.5	19.0	248	1	MABC_PONPY	Q66564 pongo pygma
28	182.5	19.0	371	1	CL46_BOVIN	Q8mh29 bos taurus
29	182.5	19.0	1268	1	PGCN_MOUSE	P55066 mus musculus
30	182.5	19.0	1268	2	Q6PIE3	Q6pie3 mus musculus
31	181.5	18.9	248	1	MABC_HYLLA	Q66554 hylobates 1

RESULT 1					ALIGNMENTS				
ID	TETN_HUMAN	STANDARD;	PRT;	202 AA.					
AC	P05452;								Q61361 mus musculus
DT	01-NOV-1988 (Rel. 09, Created)								Q80wt7 mus musculus
DT	01-APR-1993 (Rel. 25, Last sequence update)								Q8axr8 anguilla ja
DE	25-OCT-2004 (Rel. 45, Last annotation update)								Q66550 macaca fasc
DE	Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein).								Q66565 papio papio
GN	Name=TNA;								Q28518 macaca mulia
OS	Homo sapiens (Human)								P23805 bos taurus
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								P08661 rattus norv
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								O14594 homo sapien
OX	NCBI_TaxID=9606;								Q63097 rattus norv
RN	[1]								Q6z4t3 rattus norv
RP	SEQUENCE FROM N.A.								Q6z45 rattus norv
RC	TISSUE=Placenta;								Q28062 bos taurus
RX	MEDLINE=92365345; PubMed=1354271;								Q28517 macaca mulia
RA	Wewer U.M., Albrechtsen R.;								
RT	"Tetranectin, a plasminogen kringle 4-binding protein. Cloning and gene expression pattern in human colon cancer.";								
RL	Lab. Invest. 67:253-262(1992).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=92380263; PubMed=1511740; DOI=10.1016/0014-5793(92)80729-Z;								
RA	Berglund L., Petersen T.E.;								
RT	"The gene structure of tetranectin, a plasminogen binding protein.";								
RL	FEBS Lett. 309:15-19(1992).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Lung;								
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;								
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,								
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,								
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,								
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,								
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,								
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,								
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,								
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,								
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,								
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,								
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,								
RA	Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,								
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,								
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,								
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,								
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,								
RA	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.;								
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";								
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).								
RN	[4]								
RP	SEQUENCE OF 22-202, AND VARIANTS SER-55 AND MET-58.								
RX	MEDLINE=88107595; PubMed=3427041;								

RA Fuhlendorff J., Clemmensen I., Magnusson S.;  
RT "Primary structure of tetranectin, a plasminogen kringle 4 binding  
RT plasma protein: homology with asialoglycoprotein receptors and  
RT cartilage proteoglycan core protein.";  
RL Biochemistry 26:6757-6764 (1987).  
RN [5]  
RN SEQUENCE OF 1-36 FROM N.A.  
RP TISSUE=Placenta;  
RC TISSUE=Placenta;  
RC Sorensen C.B., Berglund L., Petersen T.E.;  
RT "Cloning and mapping of the murine tetranectin gene.";  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN PARTIAL SEQUENCE, MASS SPECTROMETRY, AND VARIANT GLY-106.  
RP TISSUE=Plasma;  
RC MEDLINE=20080486; PubMed=10614823;  
RA Jaquinod M., Holtet T.L., Etzerodt M., Clemmensen I., Thøgersen H.C.,  
RA Koepstorff P.;  
RT "Mass spectrometric characterisation of post-translational  
RT modification and genetic variation in human tetranectin.";  
RL Biol. Chem. 380:1307-1314 (1999).  
RN [7]  
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RP MEDLINE=97398360; PubMed=9256258; DOI=10.1016/S0014-5793(97)00664-9;  
RA Nielsen B.B., Kasstrup J.S., Rasmussen H., Holtet T.L., Graversen J.H.,  
RA Etzerodt M., Thøgersen H.C., Larsen I.K.;  
RT "Crystal structure of tetranectin, a trimeric plasminogen-binding  
RT protein with an alpha-helical coiled coil.";  
RL FEBS Lett. 412:388-396 (1997).  
RN [8]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 66-202.  
RP MEDLINE=98437604; PubMed=9757090; DOI=10.1107/S0907444997016806;  
RA Kasstrup J.S., Nielsen B.B., Rasmussen H., Holtet T.L., Graversen J.H.,  
RA Etzerodt M., Thøgersen H.C., Larsen I.K.;  
RT "Structure of the C-type lectin carbohydrate recognition domain of  
RT human tetranectin.";  
RL Acta Crystallogr. D 54:757-766 (1998).  
RN [9]  
RN -1- FUNCTION: Tetranectin binds to plasminogen and to isolated kringle  
CC 4. May be involved in the packaging of molecules destined for  
CC exocytosis.  
CC -1- SUBUNIT: Homotrimer.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Found in plasma.  
CC -1- MASS SPECTROMETRY: MW=20535.8; MW\_ERR=2.4; METHOD=Electrospray;  
CC RANGE=22-202; NOTE=Ref.6.  
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X70910; CAA50265.1; -  
DR EMBL; X70911; CAA50265.1; JOINED.  
DR EMBL; X70912; CAA50265.1; JOINED.  
DR EMBL; X64559; CAA45860.1; -  
DR EMBL; BC011024; AAH11024.1; -  
DR EMBL; X98121; CAA66803.1; -  
DR FIR; S24126; ITHUN.  
DR PDB; 1HTN; X-ray; @=-.  
DR PDB; 1TN3; X-ray; @=66-202.  
DR Genew; HGNC:11891; TNA.  
DR H-InvDB; HIX0003237; -  
DR MIM; 187520; -  
DR GO; GO:0001501; P:skeletal development; TAS.  
DR InterPro; IPR001304; Lectin C.  
DR InterPro; IPR003990; Pancreatis\_ac.  
DR Pfam; PF00059; Lectin C; 1.  
DR PRINTS; PR01504; PNCREATISAP.  
DR SMART; SM00034; CLECT. 1.  
DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.

DR PROSITE; PS0041; C\_TYPE\_LECTIN\_2; 1.  
KW 3D-structure; Direct protein sequencing; Glycoprotein; Lectin; Plasma;  
KW Polymorphism; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 202  
FT DOMAIN 77 198  
FT DISULFID 71 81  
FT DISULFID 98 197  
FT DISULFID 173 189  
FT CARBOHYD 25 25  
FT VARIANT 55 55  
FT VARIANT 58 58  
FT VARIANT 106 106  
FT STRAND 70 73  
FT STRAND 79 89  
FT HELIX 91 100  
FT TURN 101 102  
FT STRAND 104 105  
FT HELIX 111 124  
FT TURN 125 125  
FT TURN 127 128  
FT STRAND 130 140  
FT TURN 141 142  
FT STRAND 143 146  
FT TURN 147 148  
FT STRAND 151 151  
FT STRAND 157 157  
FT TURN 160 162  
FT HELIX 168 170  
FT STRAND 173 177  
FT TURN 178 182  
FT STRAND 183 187  
FT TURN 189 190  
FT STRAND 193 200  
SQ SEQUENCE 202 AA; 22567 MW; 2B0DCB5DF22E1AB8 CRC64;  
  
Query Match 100.0%; Score 960; DB 1; Length 202;  
Best Local Similarity 100.0%; Pred. No. 1.8e-77;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 EPPTQPKPKIVNAKDVVNTKMFELKSRDLTTLAQEVALLKEQQAALQTVCLKGTVMKMC 60  
DB 22 EPPTQPKPKIVNAKDVVNTKMFELKSRDLTTLAQEVALLKEQQAALQTVCLKGTVMKMC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPTGSENDALEYLRQSVGNBAEIIWGLNDMAAE 120  
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPTGSENDALEYLRQSVGNBAEIIWGLNDMAAE 141  
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGANGKWFEDKRCRDOLPYICQFI 180  
DB 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGANGKWFEDKRCRDOLPYICQFI 201  
  
QY 181 V 181  
DB 202 V 202  
  
RESULT 2  
TETN\_MOUSE  
ID TETN\_MOUSE STANDARD; PRT; 202 AA.  
AC P43025;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Tetranectin precursor (TN) (plasminogen-kringle 4 binding protein).  
GN Name=Tna;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;  
 RX MEDLINE=95137396; PubMed=7835708; DOI=10.1016/0378-1119(94)00703-U;  
 RA Soerensen C.B., Berglund L., Petersen T.E.;  
 RT "Cloning of a cDNA encoding murine tetranectin.";  
 RL Gene 152:243-245(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=96116955; PubMed=8563165;  
 RA Ibaraki K., Kozak C.A., Wewer U.M., Albrechtsen R., Young M.F.;  
 RT "Mouse tetranectin: cDNA sequence, tissue-specific expression, and  
 chromosomal mapping.";  
 RL Mamm. Genome 6:693-696(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=98072445; PubMed=9409787; DOI=10.1016/S0378-1119(97)00451-4;  
 RA Soerensen C.B., Berglund L., Petersen T.E.;  
 RT "Cloning of the murine tetranectin gene and 5'-flanking region.";  
 RL Gene 201:199-202(1997).  
 CC -!- FUNCTION: Tetranectin binds to plasminogen and to isolated kringle  
 exocytosis (By similarity).  
 CC 4. May be involved in the packaging of molecules destined for  
 -!- SUBUNIT: Homotrimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Highest expression in lung and skeletal  
 muscle.  
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X79199; CAA55791.1; -;  
 DR EMBL; U08595; AAR96811.1; -;  
 DR EMBL; X98122; CAA66804.1; -;  
 DR PIR; JC4031; JC4031.  
 DR HSP; P05452; ITN3.  
 DR MGD; MGI:104540; Tna.  
 DR GO; GO:0001501; P:skeletal development; IMP.  
 DR InterPro; IPR001304; Lectin C.  
 DR Pfam; PF00059; Lectin C; 1.  
 DR PRINTS; PR01504; PNCREATITSAP.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
 KW Lectin; Plasma; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 202 By similarity.  
 FT DOMAIN 77 198 C-type lectin.  
 FT DISULFID 71 81 By similarity.  
 FT DISULFID 98 197 By similarity.  
 FT DISULFID 173 189 By similarity.  
 FT CONFLICT 19 20 LT -> VI (in Ref. 2).  
 FT CONFLICT 84 84 A -> R (in Ref. 2).  
 FT CONFLICT 180 180 A -> R (in Ref. 2).  
 FT CONFLICT 188 188 R -> Q (in Ref. 2).  
 SQ SEQUENCE 202 AA; 22257 MW; 639E7334D5E8E04E CRC64;  
 Query Match 82.5%; Score 792; DB 1; Length 202;  
 Best Local Similarity 80.7%; Pred. No. 1.6e-62;  
 Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;  
 QY 1 EPPTOKPKIVNAKDVNTKMFELKSRDLTAEVALLKEQOALQTVCLGKTKVHKC 60  
 DB 22 ESPTPKAKKAANKKDLVSSKMFELKNRMDVLAQVALLKEQALQTVCLGKTKVHKC 81

QY 61 FLAFTQTKTFHEASEDCISRGGTSLSTPQTGSENDALYELRQSVGNENAEIWLGLNDMAAE 120  
 DB 82 LLAFATQPTTFHEASEDCISQGGTLGTQPSLENEALFEYARHSVGNNDANIWLGLNDMAAE 141  
 QY 121 GTWVDMTCARATYKNWETEITTAQPDGGKTKENCANVLGSGAANGKWFDKRCRDOLPYICQFGI 180  
 DB 142 GAWVDMTGGLLAYKNWETEITTPDGGKAENCAALSGAANGKWFDKRCRDOLPYICQFAI 201  
 QY 181 V 181  
 DB 202 V 202  
 RESULT 3  
 Q8CFZ6  
 ID Q8CFZ6 PRELIMINARY; PRT; 202 AA.  
 AC Q8CFZ6;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Tetranectin (Plasminogen binding protein).  
 GN Name=Tna;  
 OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=23382657; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC035043; AAB35043.1; -;  
 DR HSP; P05452; ITN3.  
 DR MGD; MGI:104540; Tna.  
 DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); TAS.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0005509; F:calcium ion binding; TAS.  
 DR GO; GO:0005529; F:sugar binding; TAS.  
 DR GO; GO:0001501; P:skeletal development; IMP.  
 DR InterPro; IPR001304; Lectin C.  
 DR Pfam; PF00059; Lectin C; 1.  
 DR PRINTS; PR01504; PNCREATITSAP.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
 SQ SEQUENCE 202 AA; 22255 MW; 65707A362F93A3C3 CRC64;

```
Query Match      82.5%; Score 792; DB 2; Length 202;
Best Local Similarity 80.7%; Pred. No. 1.6e-62;
Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 60
DB 22 ESPTPKAKAANAKDLVSSKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 81
QY 61 FLAFTQPKTTHASEDCISRGGLTSTPQTGSENDALVEYLRSQVGNFAETWGLNDMAAE 120
DB 82 LLFTQPKTTHASEDCISRGGLTSTPQTGSENDALVEYLRSQVGNFAETWGLNDMAAE 141
QY 121 GTWDMTGARIAYKNWETETIAQPDGKTKENC AVLGAANGKWFDRCDQLPYICQFGI 180
DB 142 GAWDMTGGLLAYKNWETETITQPDGKTKENCAALGAANGKWFDRCDQLPYICQFAI 201
QY 181 V 181
DB 202 V 202

RESULT 4
ID Q68DS3 PRELIMINARY; PRT; 133 AA.
AC Q68DS3;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Hypothetical protein DKF2p686H17246 (Fragment).
GN Name=DKF2p686H17246;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP TISSUE=Salivary gland;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Meves H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL: CR749292, CAH18147.1;
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 133 AA; 14708 MW; EFBA9A42EFE341F4 CRC64;

Query Match      75.5%; Score 725; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 9.4e-57;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VCLKGTVMKCFLAFTQPKTTHASEDCISRGGLTSTPQTGSENDALVEYLRSQVGNFA 108
DB 1 VCLKGTVMKCFLAFTQPKTTHASEDCISRGGLTSTPQTGSENDALVEYLRSQVGNFA 60
QY 109 ETWGLNDMAAGTWDMTGARIAYKNWETETIAQPDGKTKENC AVLGAANGKWFDRKRC 168
DB 61 ETWGLNDMAAGTWDMTGARIAYKNWETETIAQPDGKTKENC AVLGAANGKWFDRKRC 120
QY 169 RDQLPYICQFGIV 181
DB 121 RDQLPYICQFGIV 133

RESULT 5
Q9DDDA
ID Q9DDDA PRELIMINARY; PRT; 201 AA.
AC Q9DDDA;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Tetranectin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Wewer U.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL: AJ277116; CAC20217.1;
DR HSSP; P05452; ITN3.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
SQ SEQUENCE 201 AA; 22172 MW; 7C7F235D24426AE8 CRC64;

Query Match      75.1%; Score 720.5; DB 2; Length 201;
Best Local Similarity 73.4%; Pred. No. 3.8e-56;
Matches 130; Conservative 23; Mismatches 21; Indels 3; Gaps 1;

QY 5 QPKKIVNAKDVVNTKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 64
DB 28 QKP---AASKKGVSLKMIEDLKAMIDNISQVALLKEQALQTVCLKGTVMKCF 84
QY 65 TQTKTFHASEDCISRGGLTSTPQTGSENDALVEYLRSQVGNFAETWGLNDMAAGTW 124
DB 85 SESKTYHASEHCISQGGTLTGPQGGENDALYDMRKISGNEAEIWIGLNDMVAEGKW 144
QY 125 DMTGARIAYKNWETETIAQPDGKTKENC AVLGAANGKWFDRCDQLPYICQFGIV 181
DB 145 DMTGSPIRYKNWETETITQPDGKTKENC AVLGAANGKWFDRCDQLPYICQFGIV 201

RESULT 6
Q66KUI
ID Q66KUI PRELIMINARY; PRT; 193 AA.
AC Q66KUI;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE MGC85438 protein.
GN Name=MGC85438;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RL "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL initiative."
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL initiative."
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC078559; AAH78559.1; -;  
 DR InterPro: IPR001304; Lectin\_C.  
 DR Pfam: PF00059; Lectin\_C; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
 SQ SEQUENCE 193 AA; 22097 MW; 2FD58EF734E626D7 CRC64;

Query Match 47.4%; Score 455; DB 2; Length 193;  
 Best Local Similarity 49.1%; Pred. No. 1.7e-32;  
 Matches 86; Conservative 34; Mismatches 49; Indels 6; Gaps 2;

QY 4 TQPKKIVNAKDVNTKMFELKSLRLDTLAEVALLKEQALQTVCLKGTVMKCFLA 63  
 DB 22 TUSAK--LTKQDHRKSEKQGLQTDKWLREINSLKEMQALQTVCLKGTVMKCF 79  
 QY 64 FTQTKTFHEASDCISRGGLTSTPQTSNDALYELRSGVNEAIEIWLGLNDMAAEGTW 123  
 DB 80 FEETKHFHEANEDCIAKGGTLAI PRDSENNALRDYKSLHSGEFLGINDMVEGKF 139  
 QY 124 VDMTGARIYKWEITEAOPGGKTENCANVLSGAANGKWFDRCDQLPYICQF 178  
 DB 140 VDVNGVAITYFWNE----RIPKGRKKNALLNQASQGWVDEVCRSLKYYICEF 190

RESULT 7  
 Q28008 ID Q28008 PRELIMINARY; PRT; 197 AA.  
 AC Q28008;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE C-type lectin homolog precursor.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cartilage;  
 RA Neame P.J., Boynton R.E.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U22298; AAC18614.1; -;  
 DR HSSP; P05452; ITN3.  
 DR GO; GO:0005529; F:sugar binding; IEA.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR Pfam; PF00059; Lectin\_C; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.

KW Lectin; Signal.  
 FT SIGNAL 1 24 Potential.  
 FT CHAIN 25 197 C-type lectin homolog.  
 SQ SEQUENCE 197 AA; 22215 MW; AAAC4280F41ACOF4 CRC64;

Query Match 44.9%; Score 431.5; DB 2; Length 197;  
 Best Local Similarity 51.9%; Pred. No. 2.1e-30;  
 Matches 81; Conservative 29; Mismatches 43; Indels 3; Gaps 1;

QY 25 ELKSLRLDTLAEVALLKEQALQTVCLKGTVMKCFLAFTQTKTFHEASDCISRGGLT 84  
 DB 43 DLKTOVEKLREVNALKEMQALQTVCLRGTRFKKCYLAEEGLKHFHEANEDCISKGGLT 102  
 QY 85 STPQTSNDALYELRSGVNEAIEIWLGLNDMAAEGTWDMTGARIYKWEITEAOP 144  
 DB 103 VVPSADSEINALRDYKSLPGVNDFWLGINDMVAEGKFDVINGLAISFLNWD---QAOP 159  
 QY 145 DGGKTENCANVLSGAANGKWFDRCDQLPYICQF 180  
 DB 160 NGGKRENCALFSQSAQKWSDEACHSKRYICEFTI 195

RESULT 8  
 Q28P4 ID Q28P4 PRELIMINARY; PRT; 196 AA.  
 AC Q28P4;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE C-type lectin superfamily 1.  
 GN Name=Clectsl;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ES129/SV;  
 RA Neame P.J., Grimm D.R.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF317204; AAG48620.1; -;  
 DR HSSP; P05452; IHTN.  
 DR MGD; MGI:2685642; Clectsl.  
 DR GO; GO:0005529; F:sugar binding; IEA.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR Pfam; PF00059; Lectin\_C; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
 KW Lectin.  
 SQ SEQUENCE 196 AA; 22191 MW; 9AB4C809D119B852 CRC64;

Query Match 44.8%; Score 430.5; DB 2; Length 196;  
 Best Local Similarity 48.5%; Pred. No. 2.6e-30;  
 Matches 83; Conservative 30; Mismatches 47; Indels 11; Gaps 2;

QY 8 KKVNAKDVNTKMFELKSLRLDTLAEVALLKEQALQTVCLKGTVMKCFLAFTQ 67  
 DB 34 KRVKAKDD-----DLKSQVEKLREVNALKEMQALQTVCLRGTRFKKCYLAEEGL 85  
 QY 68 KTFHEASDCISRGGLTSTPQTSNDALYELRSGVNEAIEIWLGLNDMAAEGTWDMT 127  
 DB 86 KHYTHEANEDCISKGGLTVVPRNSDEINALRDYKSLPGVNDFWLGINDMVTEGFLDVH 145  
 QY 128 GARIYKWEITEAOPGGKTENCANVLSGAANGKWFDRCDQLPYICQF 178  
 DB 146 GFVAVSFLNWD---RAQPSGKRENCALFSQSAQKWSDEACHSKRYICEF 193

RESULT 9  
 CLEI\_HUMAN ID CLEI\_HUMAN STANDARD; PRT; 197 AA.  
 AC 075596;



```

DR HSP; P05452; 1HTN.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lectin C; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Direct protein sequencing; Lectin.
FT DOMAIN 43 161 C-type lectin.
FT DISULFID 37 47
FT DISULFID 64 160
FT DISULFID 136 152
SQ SEQUENCE 166 AA; 18432 MW; 53BF812DEA5C6119 CRC64;

Query Match 42.5%; Score 408; DB 1; Length 166;
Best Local Similarity 44.5%; Pred. No. 2.1e-28;
Matches 77; Conservative 31; Mismatches 53; Indels 12; Gaps 2;

QY 6 KPKIVNAKDVVNTKMPPEELKSLRLDTLAQEVALLKEQALQTVCLKGTQVHKMCKFLAFT 65
DB 2 KPSKSGKGD-----LNREIDKLWREVNSLKEMQALQTVCLKGTQVHKMCKFLAFT 52
QY 66 QTKTFHEASEDCISRGGLTSTPTQSGENDALYEYLRQSVGNBAEIVLGLNDMAAEGTWVD 125
DB 53 GSKSYHAANEDCIAQGGTSLIPRSSDEGNSLSYAKSLVGARDFWIGVNDWTTGSKFVD 112
QY 126 MTGARIYKNTWTETTAQPDGKNTENCVLGSAANGKWFKRCRDLPIYICOF 178
DB 113 VNLPLITTFNWD---RSKPVGGTRENCAVASTSGQKNSDDVCRSEKRYICEY 162

RESULT 12
SCGF RAT STANDARD; PRT; 328 AA.
AC 088201;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Stem cell growth factor precursor (Lymphocyte secreted C-type lectin).
GN Name=Scgf;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone;
RX MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbr.1998.9073;
RA Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA Sekine S., Hiraoka A.;
RT "Isolation and characterization of a cDNA for human, mouse, and rat
RT full-length stem cell growth factor, a new member of C-type lectin
RT superfamily."
RL Biochem. Biophys. Res. Commun. 249:124-130(1998).
CC -!- FUNCTION: Stimulates the proliferation and differentiation of
CC hematopoietic precursor cells from various lineages, including
CC erythrocytes, lymphocytes, granulocytes and macrophages. Acts
CC synergistically with other cytokines, including IL-3, G-CSF, GM-CSF
CC and FLT3 ligand. Suppresses SCF-stimulated erythrocyte
CC proliferation (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted (By similarity).
CC -!- PTM: O-glycosylated. Probably sulfated on the O-glycans (By
CC similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009246; BAA32406.1; -
DR HSP; P05452; 1HTN.

```

---

```

DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0008083; P:growth factor activity; IDA.
DR GO; GO:0005529; P:sugar binding; NAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00059; Lectin C; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Glycoprotein; Growth factor; Lectin; Signal.
FT SIGNAL 1 21 By similarity.
FT CHAIN 22 328 Stem cell growth factor.
FT DOMAIN 188 325 C-type lectin.
SQ SEQUENCE 328 AA; 36387 MW; C115188AC23D6257 CRC64;

Query Match 23.2%; Score 222.5; DB 1; Length 328;
Best Local Similarity 28.6%; Pred. No. 1.6e-11;
Matches 55; Conservative 33; Mismatches 73; Indels 31; Gaps 5;

QY 17 VVNTKMF-----ELKSLRLDTLAQEVALLKEQALQTV-----CLKGTQVHKMCKFLA 63
DB 136 VLDTRVVELTQGLRLRLDAASDTRDSVQALKEVQVRSQEHGRLEGCLKGLRGLHKCFLL 195
QY 64 FTQTKTFHEASEDCISRGGLTSTPTQSGENDALYEYLRQSVGNBAEIVLGLNDMAAEGT 122
DB 196 SRDFETQAAQAQKCKARGGSLAQPADRQMDALSRYLAALAPYNWVWLVGHVDRRSEGL 255
QY 123 VWDMTGARIYKNTWTETTA-----QPDGKNTENCVLGSAANGKWFK 166
DB 256 YLFENGQVRSFFAMHRAUSPESGAQPSAASHPLSPDQNGGILENC-VAQASDDGSWWDH 314
QY 167 RCRDLPIYICOF 178
DB 315 DCERLYFVCFE 326

RESULT 13
SCGF MOUSE STANDARD; PRT; 328 AA.
AC 088200; OBC946; Q9CTF0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Stem cell growth factor precursor (Lymphocyte secreted C-type lectin).
GN Name=Scgf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Calvaria;
RX MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbr.1998.9073;
RA Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA Sekine S., Hiraoka A.;
RT "Isolation and characterization of a cDNA for human, mouse, and rat
RT full-length stem cell growth factor, a new member of C-type lectin
RT superfamily."
RL Biochem. Biophys. Res. Commun. 249:124-130(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```



Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Raheij J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences".  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [3]  
 SEQUENCE OF 114-328 FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Embryo;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Otsu N., Saito R., Suzuki H., Yamana H., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Glissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,  
 RA Petrovsky N., Pillai K., Pontius J.R., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed J.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wu L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs".  
 RL Nature 420:563-573(2002).  
 CC -!- FUNCTION: Stimulates the proliferation and differentiation of  
 hematopoietic precursor cells from various lineages, including  
 erythrocytes, lymphocytes, granulocytes and macrophages. Acts  
 synergistically with other cytokines, including IL-3, G-CSF, GM-CSF  
 and FLT3 ligand. Suppresses SCF-stimulated erythrocyte  
 proliferation (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted (By similarity).  
 CC -!- PTM: O-glycosylated. Probably sulfated on the O-glycans (By  
 similarity). Contains 1 C-type lectin family domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AB09245; BAA32405.1;  
 CC EMBL: BC020001; AH020001.3;  
 CC EMBL: AK042963; BAC31421.1;  
 CC EMBL: AK003813; BAB23010.1;  
 CC HSSP: P05452; 1HTN.  
 CC MGI: 1298219; Scgf.  
 CC GO: GO:0005576; C:extracellular; ISS.  
 CC GO: GO:0008083; F:growth factor activity; ISS.  
 CC GO: GO:0005529; F:sugar binding; NAS.  
 CC GO: GO:0008284; P:positive regulation of cell proliferation; ISS.

DR InterPro: IPR001304; Lectin C.  
 DR InterPro: IPR008994; Nucleic\_acid\_OB.  
 DR Pfam: PF00059; Lectin\_C; 1.  
 DR SMART: SM00034; CLECT; 1.  
 DR PROSITE: PS00615; C-TYPE LECTIN 1; 1.  
 DR PROSITE: PS00041; C-TYPE LECTIN 2; 1.  
 KW Glycoprotein; Growth factor; Lectin; Signal.  
 FT SIGNAL 1 21 By similarity.  
 FT CHAIN 22 328 Stem cell growth factor.  
 FT DOMAIN 188 325 C-type lectin.  
 FT CONFLICT 114 117 TVIL -> FFTV (in Ref. 3).  
 SQ SEQUENCE 328 AA; 36451 MW; 309C17A861EE135C CRC64;  
 Query Match 23.1%; Score 221.5; DB 1; Length 328;  
 Best Local Similarity 27.8%; Pred. No. 26-11; 76; Indels 31; Gaps 5;  
 Matches 55; Conservative 36; Mismatches 31;  
 QY 11 VNAKDVVNTKMF-----ELKSLDTLAEVALLKEQ-----ALQTVCLKGTQVH 57  
 DB 130 LHVRLHVLDTRVVELTQGLRQLRDAASDTRDSVQALKEVQDRAEQEHGRLEGLKGLRLG 189  
 QY 58 MKCFLAFTQTKTFHASEDCISRGSTLSTPQTGSENDALYEYLRQSVG-NEAEIWLGLND 116  
 DB 190 HKCFLLSRDFETQAAQAQARCKARGSLAQAPADQOQMDALSRYLRAALAPYNPVWVGLGVD 249  
 QY 117 MAAGTVMVTGARTAYKNWETETFA-----OPDGGKTENCVLGSAAN 160  
 DB 250 RSEGLYFENGQVRVSFAWIRAFSLGSAFSAATHPLSPDQNGGVLENC-VAQASDD 308  
 QY 161 GKWFDRCRDQLPYICQF 178  
 DB 309 GSWBHDCCERRLYFVCEP 326  
 RESULT 14  
 ID SCGF HUMAN STANDARD; PRT; 323 AA.  
 AC Q9Y240;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Stem cell growth factor precursor (Lymphocyte secreted C-type lectin)  
 GN (p47).  
 OS Homo sapiens (Human);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 22-39; 229-243 AND 314-323, TISSUE  
 RP SPECIFICITY, AND O-GLYCOSYLATION.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=98113146; PubMed=9442024; DOI=10.1074/jbc.273.4.1911;  
 RA Banwarth S., Giordanengo V., Lesimple J., Lefebvre J.-C.;  
 RT "Molecular cloning of a new secreted sulfated mucin-like protein with  
 a C-type lectin domain that is expressed in lymphoblastic cells";  
 RL J. Biol. Chem. 273:1911-1916(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbrc.1998.9073;  
 RA Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,  
 RA Sekine S., Hiraoka A.;  
 RT "Isolation and characterization of a cDNA for human, mouse, and rat  
 full-length stem cell growth factor, a new member of C-type lectin  
 superfamily";  
 RL Biochem. Biophys. Res. Commun. 249:124-130(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99216429; PubMed=10198175; DOI=10.1006/geno.1999.5762;  
 RA Banwarth S., Giordanengo V., Grosgeorge J., Turc-Carel C.,  
 RA Lefebvre J.-C.;  
 RT "Cloning, mapping, and genomic organization of the LSLCL gene,



encoding a new lymphocytic secreted mucin-like protein with a C-type lectin domain: a new model of exon shuffling.";  
Genomics 57:316-317(1999).

[4]

SEQUENCE FROM N.A.

TISSUE-Skin;

MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

STRAUSBERG R.B., FEINGOLD E.A., GROUSE L.H., DERGE J.G.,  
KLAUSNER R.D., COLLINS F.S., WAGNER L., SHENMEN C.M., SCHULER G.D.,  
ALTSCHUL S.F., ZEEBORG B., BETOW K.H., SCHAEFER C.F., BHAT N.K.,  
HOPKINS R.F., JORDAN H., MOORE T., MAX S.I., WANG J., HSIEH F.,  
DACHENKO L., MARUSHINA K., FARMER A.A., RUBIN G.M., HONG L.,  
STAPLETON M., SOARES M.B., RONALDO M.F., CASAVANT T.L., SCHEETZ T.E.,  
BROWNSTEIN M.J., USLDIN T.B., TOSHIYUKI S., CARLINCI P., PRANGE C.,  
RAHA S.S., LOQUELLANO N.A., PETERS G.J., ABRAMSON R.D., MULLAHDY S.J.,  
BOSAK S.A., MCEWAN P.J., MCKERNAN K.J., MALEK J.A., GUNARATNE P.H.,  
RICHARDS S., WORLEY K.C., HALE S., GARCIA A.M., GAY L.J., HULYK S.W.,  
VILLALON D.K., MUZYNY D.M., SODERGREN E.J., LU X., GIBBS R.A.,  
FAHEY J., HELTON E., KETEMAN M., MADAN A., RODRIGUES S., SANCHEZ A.,  
WHITING M., MADAN A., YOUNG A.C., SHEVCHENKO Y., BOUFFARD G.G.,  
BLAKESLEY R.W., TOUCHMAN J.W., GREEN E.D., DICKSON M.C.,  
RODRIGUEZ A.C., GRIMWOOD J., SCHMUTZ J., MYERS R.M.,  
BUTTERFIELD V.S.N., KRZYWINSKI M.I., SKALEKA U., SMAILUS D.E.,  
SCHNERCH A., SCHEIN J.E., JONES S.J.M., MARRA M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[5]

FUNCTION, AND TISSUE SPECIFICITY.

MEDLINE=21917201; PubMed=11920266; DOI=10.1038/sj/thj/6200118;

HIRAKO A., YANO K.-I., KAGAMI N., TAKESHIGE K., MIO H., ANAZAWA H.,  
SUGIMOTO S.;  
"Stem cell growth factor: in situ hybridization analysis on the gene  
expression, molecular characterization and in vitro proliferative  
activity of a recombinant preparation on primitive hematopoietic  
progenitor cells.";  
Hematol. J. 2:307-315(2001).

[6]

SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.

MEDLINE=21663069; PubMed=11803813; DOI=10.1016/S0764-4469(01)01392-0;  
PERRIN C., BAYLE J., BANNWARTH S., MICHELIS J.-F., HEUDIER P.,  
LEFEBVRE J.-C., GIORDANENGO V.;  
"Expression of LSuCL, a new C-type lectin, is closely restricted, in  
bone marrow, to immature neutrophils.";  
C. R. Acad. Sci. III, Sci. Vie 324:1125-1132(2001).  
"-!- FUNCTION: Stimulates the proliferation and differentiation of  
hematopoietic precursor cells from various lineages, including  
erythrocytes, lymphocytes, granulocytes and macrophages. Acts  
synergistically with other cytokines, including IL-3, G-CSF, GM-CSF  
and FLT3 ligand. Suppresses SCF-stimulated erythrocyte  
proliferation.

"-!- SUBCELLULAR LOCATION: Cytoplasmic and secreted.

"-!- TISSUE SPECIFICITY: Expressed in skeletal tissues including bone  
marrow, chondrocytes, primary ossification center-associated  
cells, the perichondrium and periosteum. Lower levels of  
expression were detected in spleen, thymus, appendix and fetal  
liver.

"-!- DEVELOPMENTAL STAGE: In the bone marrow, expression is limited to  
immature neutrophils. Expression was not detected in circulating  
mature neutrophils.

"-!- PTM: O-glycosylated. Probably sulfated on the O-glycans.

"-!- SIMILARITY: Contains 1 C-type lectin family domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF020044; AAC39569.1; -.

EMBL; AB009244; BAA32404.1; -.

DR

DR EMBL; AF087658; AAD26533.1; -.  
DR EMBL; BC005810; AAH05810.1; -.  
DR HSSP; P05452; IHTN.  
DR Genew; HGNC:10576; SCGF.  
DR MIM; 604713; -.  
DR GO; GO:0005576; C:extracellular; IDA.  
DR GO; GO:0008083; F:growth factor activity; IDA.  
DR GO; GO:0005529; F:sugar binding; NAS.  
DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.  
DR InterPro; IPR001304; Lectin C.  
DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR Pfam; PF00059; Lectin C; 1.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_2; 1.  
KW Direct protein sequencing; Glycoprotein; Growth factor; Lectin;  
KW SIGNAL.  
FT CHAIN 1 21 Stem cell growth factor.  
FT DOMAIN 22 323 C-type lectin.  
FT SITE 183 320 Cell attachment site (Potential).  
FT SITE 61 63  
SQ SEQUENCE 323 AA; 35694 MW; D13604CDAF087427 CRC64;  
Query Match 23.0%; Score 220.5; DB 1; Length 323;  
Best Local Similarity 29.5%; Pred. No. 2.4e-11; Mismatches 37; Indels 29; Gaps 7;  
Matches 57; Conservative 37; Mismatches 70; Indels 29; Gaps 7;  
QY 4 TQPKKIVNAKKDVVN-TKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTGVHMKCFL 62  
DB 140 TQGLRQLNNAAGTRDAVQALQEAQGRAE-----REHGRLEG-CLKGLRLGHKCF 189  
QY 63 APTQTKTFHEASEDCISRGGLSTPTQSGENDALYEYLRQSVG-NEAEIWLGLNDMAAEG 121  
DB 190 LSRDFEAAAQAARCTARGGSLAQPADRQMEALTRYLRAALAPYNWFWLGVHDDRAEG 249  
QY 122 TWDMTGARIAYKNW---ETEITA-----OPDGKTEKNCVLGSAANGKWF 165  
DB 250 LYLFGNGRVSVFFAWHRSFPRPELGAQPSASPHLSPDQFNGTLENC-VAQASDDGSWMD 308  
QY 166 KRCRDQLPYICOF 178  
DB 309 HDCQRLYYVCEF 321  
RESULT 15  
PSPD BOVIN  
ID PSPD BOVIN STANDARD; PRT; 369 AA.  
AC P35246;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).  
GN Name=SFTPD; Synonyms=SFTP4;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]\_  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 208-247.  
RC TISSUE=Lung;  
RA MEDLINE=93170856; PubMed=8436402;  
RA Lim B.L., Lu J., Reid K.B.M.;  
RT "Structural similarity between bovine conglutinin and bovine lung  
RT surfactant protein D and demonstration of liver as a site of synthesis  
RT of conglutinin.";  
RL Immunology 78:159-165(1993).  
CC "-!- FUNCTION: Contributes to the lung's defense against inhaled  
CC microorganisms. Binds strongly maltose residues and to a lesser  
CC extent other alpha-glucosyl moieties. It could participate in the  
CC extracellular reorganization or turnover of pulmonary surfactant.  
CC "-!- SUBUNIT: Oligomeric complex of 4 set of homotrimers.  
CC "-!- SUBCELLULAR LOCATION: Extracellular.

```

CC CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated protein: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC CC -1- SIMILARITY: Contains 1 collagenous domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X75911; CAA53510.1; -.
CC DR HSSP; P35247; 1B08.
CC DR GO; GO:0030139; C:endocytic vesicle; ISS.
CC DR GO; GO:0005576; C:extracellular; ISS.
CC DR GO; GO:0005764; C:lysosome; ISS.
CC DR GO; GO:0008367; F:bacterial binding; ISS.
CC DR GO; GO:0005529; F:sugar binding; ISS.
CC DR GO; GO:0048286; F:palveolus development; ISS.
CC DR GO; GO:0045087; P:innate immune response; ISS.
CC DR GO; GO:0048246; P:macrophage chemotaxis; ISS.
CC DR GO; GO:0045085; P:negative regulation of interleukin-2 biosyn. ; ISS.
CC DR GO; GO:0042130; P:negative regulation of T-cell proliferation; ISS.
CC DR GO; GO:0006800; P:positive regulation of phagocytosis; ISS.
CC DR GO; GO:0050766; P:positive regulation of oxygen species metabolism; ISS.
CC DR GO; GO:0006998; P:positive regulation of phagocytosis; ISS.
CC DR GO; GO:0001817; P:receptor mediated endocytosis; ISS.
CC DR InterPro; IPR008160; Collagen.
CC DR InterPro; IPR001304; Collagen.
CC DR Pfam; PF01391; Collagen; 2.
CC DR Pfam; PF00059; Lectin_C; 1.
CC DR ProDom; PD000007; Clg_helix; 1.
CC DR SMART; SM00034; CUECT; 1.
CC DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
CC DR PROSITE; PS00041; C TYPE LECTIN 2; 1.
CC KW Calcium; Coiled coil; Collagen; Direct protein sequencing;
CC KW Gaseous exchange; Glycoprotein; Hydroxylation; Lectin; Repeat; Signal;
CC KW Surface film.
CC FT SIGNAL 1 20 By similarity.
CC FT CHAIN 21 369 Pulmonary surfactant-associated protein
CC FT FT D.
CC FT DOMAIN 46 216 Collagen-like.
CC FT DOMAIN 217 248 Coiled coil (Potential).
CC FT DOMAIN 254 369 C-type lectin.
CC FT DISULFID 275 367 By similarity.
CC FT DISULFID 345 359 By similarity.
CC FT CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
CC FT MOD_RES 78 78 Hydroxyproline (By similarity).
CC FT MOD_RES 87 87 5-hydroxylysine (By similarity).
CC FT MOD_RES 95 96 Hydroxyproline (By similarity).
CC FT MOD_RES 99 99 5-hydroxylysine (By similarity).
CC FT MOD_RES 165 165 Hydroxyproline (By similarity).
CC FT MOD_RES 171 171 Hydroxyproline (By similarity).
CC FT SEQUENCE 369 AA; 373361 MW; 07D88B324E0AEB2E3C64;
CC
CC Query Match 21.6%; Score 207.5; DB 1; Length 369;
CC Best Local Similarity 33.1%; Pred.No. 4e-10;
CC Matches 54; Conservative 27; Mismatches 61; Indels 21; Gaps 7;
CC
CC QY 28 SRLDTLAEVALLKEQ-QALQTVCLK-----GTKVHKCFLAFTQKTFHEASEDC 77
CC DB 216 AEVNALRQRVGILEGQLRQNAFSQYKKAMLPFNGRSVGKIKFTVGSEKTFQDAQQIC 275
CC QY 78 ISRGGTLSPQTGSENDALVEYLRSQVGNAEIWLGLNDWAEGTWDMTGARIAYKNWE 137
CC DB 276 TQAGGQLPSRSGAENEATQL---ATAQKAFLMSDTRKEGTFIYPTGPELVYSNWA 332
CC QY 138 TEITAPQ--DGGKTENC AVLSGAANGKWFDRCRDLPYICQF 178

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 39 Seconds  
(without alignments)  
446.544 Million cell updates/sec

Title: US-09-445-576A-7

Perfect score: 960

Sequence: 1 EPPTQPKKIVNAKDVNT.....KWFDRKRDQLPYICQFGIV 181

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	202	1 TTHUN	tetranectin precu
2	792	82.5	202	2 JC4031	tetranectin precu
3	408	42.5	166	1 A37289	tetranectin homo
4	207.5	21.6	369	2 S33603	surfactant protein
5	186.5	19.4	244	1 LNM5MC	mannose-binding le
6	186.5	19.4	1257	2 S28764	neurocan precursor
7	185.5	19.3	248	1 LNHUMC	mannose-binding le
8	182.5	19.0	1268	2 S52781	neurocan - mouse
9	180	18.8	883	2 S57653	brevican precursor
10	178.5	18.6	371	1 JN0450	conglutinin precu
11	178.5	18.6	371	2 I45878	conglutinin - bovi
12	177.5	18.5	244	1 LNR1MC	mannose-binding le
13	177	18.4	309	1 S34198	IgE Fc receptor II
14	177	18.4	912	2 A54423	brevican precursor
15	176.5	18.4	374	1 A42046	surfactant protein
16	176	18.2	301	2 A53570	collectin-43 - bov
17	175	18.2	330	2 T46256	brevican - human (
18	172.5	18.0	375	1 A45225	pulmonary surfacta
19	172	17.9	883	2 A49126	brevican precursor
20	171.5	17.9	238	1 LNR1MA	mannose-binding le
21	166.5	17.3	239	1 LNM5MA	mannose-binding le
22	166	17.3	331	1 LNM5ER	IgE Fc receptor, l
23	163	17.0	1456	1 A36563	mannose receptor p
24	160	16.7	1340	2 A39808	proteoglycan core
25	160	16.7	1479	2 T42710	mannose receptor,
26	160	16.7	2327	2 T42630	aggreccan - bovine
27	157.5	16.4	175	2 A37194	pancreatic thread
28	156.5	16.3	321	2 T26152	hypothetical prote
29	152	15.8	174	2 S54979	pancreatitis-assoc

ALIGNMENTS

RESULT 1

TTHUN

tetranectin precursor [validated] - human

N:Alternate names: plasminogen-kringle 4 binding protein

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1992 #sequence revision 03-Aug-1995 #text\_change 09-Jul-2004

C:Accession: S24126; A56835; A29747; I38359; S19865

R:Berglund, L.; Petersen, T.E.

FEBS Lett. 309, 15-19, 1992

A>Title: The gene structure of tetranectin, a plasminogen binding protein.

A:Reference number: S24126; MUID:92380263; PMID:1511740

A:Accession: S24126

A:Molecule type: DNA

A:Residues: 1-202 <BER>

A:Cross-References: UNIPROT:P05452; EMBL:X70911

R:Wewer, U.M.; Albrechtsen, R.

Lab. Invest. 67, 253-262, 1992

A>Title: Tetranectin, a plasminogen kringle 4-binding protein. Cloning and gene expressi

A:Reference number: A56835; MUID:92365345; PMID:1354271

A:Accession: A56835

A:Molecule type: mRNA

A:Residues: 1-202 <WEW>

A:Cross-References: EMBL:X64559; NID:g37408; PIDN:CAA45860.1; PID:g37409

A:Experimental source: Placenta

R:Fuhlendorff, J.; Clemmensen, I.; Magnusson, S.

Biochemistry 26, 6757-6764, 1987

A>Title: Primary structure of tetranectin, a plasminogen kringle 4 binding plasma protein

A:Reference number: A29747; MUID:86107595; PMID:3427041

A:Accession: A29747

A:Molecule type: protein

A:Residues: 22-105, 'G', 107-202 <FUH>

A:Experimental source: plasma

A>Note: S5-Ser and 58-Met were also found

C:Genetics:

A:Gene: GDB:TNA

A:Cross-References: GDB:135032; OMIM:187520

A:Map position: 3p22-3p21.3

A:Introns: 37/1, 70/1

C:Complex: homotetramer

C:Superfamily: tetranectin; C-type lectin homology

C:Keywords: glycoprotein; plasma; tetramer

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-202/Product: tetranectin #status experimental <MAT>

F:71-197/Domain: C-type lectin homology <LCH>

F:25/Binding site: carbohydrate (Thr) (covalent) #status experimental

F:71-81, 98-197, 173-189/Disulfide bonds: #status experimental

Query Match 100.0%; Score 960; DB 1; Length 202;

Best Local Similarity 100.0%; Pred. No. 4e-78;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFBEELKRLDLTAQEVALLKEQQAQTVCCLKGVHMKC 60

Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTQVHMKC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
Db 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180  
Db 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 201  
QY 181 V 181  
Db 202 V 202

## RESULT 2

JC4031  
tetranectin precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: JC4031  
R:Sorensen, C.B.; Berglund, L.; Petersen, T.E.  
Gene 152, 243-245, 1995  
A:Title: Cloning of a cDNA encoding murine tetranectin.  
A:Reference number: JC4031; MUID:95137396; PMID:7835708  
A:Accession: JC4031  
A:Molecule type: mRNA  
A:Residues: 1-202 <SOR>  
A:Cross-references: UNIPROT:P43025; EMBL:X79199; MID:g671561; PIDN:CAA55791.1; PID:g6715  
A:Experimental source: lung  
C:Comment: This protein binds plasminogen, and may play a role in invasive cancer.  
F:1-21/Domain: tetranectin; C-type lectin homology  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-202/Product: tetranectin #status predicted <WAT>  
F:71-197/Domain: C-type lectin homology <LCH>

Query Match 82.5%; Score 792; DB 2; Length 202;  
Best Local Similarity 80.7%; Pred. No. 3.6e-63;  
Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;  
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTQVHMKC 60  
Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTQVHMKC 81  
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120  
Db 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141  
QY 121 GTWVDMTGARIAYKNWETEITAPQDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180  
Db 142 GTWVDMTGARIAYKNWETEITAPQDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 201  
QY 181 V 181  
Db 202 V 202

## RESULT 3

A37289  
tetranectin homolog - reef shark  
C:Species: Carcharhinus springeri (reef shark)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: A37289; A37287  
R:Neame, P.J.; Young, C.N.; Treep, J.T.  
Protein Sci. 1, 161-168, 1992  
A:Title: Primary structure of a protein isolated from reef shark Carcharhinus springeri  
A:Reference number: A37289; MUID:93284081; PMID:1304877  
A:Accession: A37289  
A:Molecule type: protein  
A:Residues: 1-166 <NE2>  
A:Cross-references: UNIPROT:P26258  
C:Superfamily: tetranectin; C-type lectin homology

C:Keywords: cartilage  
F:37-160/Domain: C-type lectin homology <LCH>  
F:37-47,64-160,136-152/Diulfide bonds: #status predicted

Query Match 42.5%; Score 408; DB 1; Length 166;  
Best Local Similarity 44.5%; Pred. No. 4.3e-29;  
Matches 77; Conservative 31; Mismatches 53; Indels 12; Gaps 2;  
QY 6 KFKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTQVHMKCFLAFT 65  
Db 2 KFKSGKSGKDD-----LRNEIDKLWREVNLSKENQALQTVCLKGTQVHMKCFLAFT 52  
QY 66 QTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAEAGTWD 125  
Db 53 GSKSYHAANEDCIAQGGTLSTPSSDEGNSLSYAKSLVGARDFWIGVNDMTTEGKFDV 112  
QY 126 MTGARIAYKNWETEITAPQDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQF 178  
Db 113 VNGLPITVFNWD---RSKPVGGTRENCAVASTSGGKWSDDVCSEKRYICEY 162

## RESULT 4

S33603  
surfactant protein D - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 02-Dec-1993 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S33603  
R:Lim, B.L.; Lu, J.; Reid, K.B.M.  
Immunology 78, 159-165, 1993  
A:Title: Structural similarity between bovine conglutinin and bovine lung surfactant prot  
A:Reference number: S33603; MUID:93170856; PMID:8436402  
A:Accession: S33603  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-369 <LIM>  
C:Superfamily: pulmonary surfactant protein D; C-type lectin homology  
F:248-367/Domain: C-type lectin homology <LCH>

Query Match 21.6%; Score 207.5; DB 2; Length 369;  
Best Local Similarity 33.1%; Pred. No. 7.6e-11;  
Matches 54; Conservative 27; Mismatches 61; Indels 21; Gaps 7;  
QY 28 SRLTLAQEVALLEKEQ-QALQTVCLK-----GPKVHMKFLAFTQTKTFHEASEDC 77  
Db 216 AEVALQRVGLLEGQLQRLQNAFQYKQKAMLPNGRSVGEKIFKTVGSEKTFQDAQOIC 275  
QY 78 ISRGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAEAGTWDMTGARIAYKNWE 137  
Db 276 TOAGGQLPSRSGAENEALTQ---ATAQNKAAPFLMSDTRKEGTFIYPTGPELAVSNWA 332  
QY 138 TEITAAP--DGGKTENCACVLSGAANGKWFDRKCRDQLPYICQF 178  
Db 333 PQ---EPNDDGG-SENCVEI--FPNGKWNKDKVCGEQLRVICEF 369

## RESULT 5

LNMSMC  
mannose-binding lectin C precursor - mouse  
N:Alternate names: Ra-reactive factor P28a  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Jun-1993 #sequence\_revision 20-Feb-1998 #text\_change 09-Jul-2004  
C:Accession: I48651; B46466; A42574; C42574  
R:Sastri, R.; Wang, J.S.; Brown, D.C.; Ezekowitz, R.A.; Tauber, A.I.; Sastri, K.N.  
Mamm. Genome 6, 103-110, 1995  
A:Title: Characterization of murine mannose-binding protein genes Mb11 and Mb12 reveals f  
A:Reference number: I48650; MUID:95284466; PMID:7766991  
A:Accession: I48651  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-244 <SAS1>  
A:Cross-references: UNIPROT:P41317; EMBL:U09016; MID:g773286; PIDN:AAA82010.1; PID:g77328  
R:Sastri, K.; Zahedi, K.; Lellias, J.M.; Whitehead, A.S.; Ezekowitz, R.A.  
J. Immunol. 147, 692-697, 1991

F;23-1257/Product: neurocan #status predicted <MAT>  
F;176-253/Domain: link protein repeat homology <LNK1>  
F;274-355/Domain: link protein repeat homology <LNK2>  
F;364-366/Region: cell attachment (R-G-D) motif  
F;953-984/Domain: EGF homology <EGF>  
F;1029-1149/Domain: C-type lectin homology <LCH>  
F;1156-1212/Domain: complement factor H repeat homology <FHD>  
F;121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted  
F;944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

Query Match 19.4%; Score 186.5; DB 2; Length 1257;  
Best Local Similarity 32.7%; Pred. No. 2.3e-08;  
Matches 48; Conservative 20; Mismatches 52; Indels 27; Gaps 8;

Qy 49 VCLAGT-----KVHMKCFLAFTQTKFHEASEDCISRGGTSTPQTGSENDALVEYL 100  
Db 1021 LCEKDTGECDRGWKFGQHCYVYFAHRAWEDRCRRAGHLTSVHSPEEHKFI---- 1076  
Qy 101 RQSVGNENAEIWLGLNDMAAEG--TWVDMTGARIYKNWETETITAPQD-----GGKTENCAV 154  
Db 1077 NSFQHE-NSWIGLNDRTVERDFQWTDNTG--LQYENWREK-----QPDNPFAGG--EDCVV 1127  
Qy 155 LSGAANGKWFDKRCRDQLPYTCQFGIV 181  
Db 1128 MVAHENGKRWNDVPCNYNLPYVCKGTV 1154

RESULT 7

LNHUMC  
mannose-binding lectin precursor [validated] - human  
N/Alternate names: mannan-binding protein  
C/Species: Homo sapiens (man)  
C/Date: 31-Mar-1989 #sequence revision 30-Sep-1991 #text change 09-Jul-2004  
C/Accession: J01115; S05641; A34978; J00027; JX0319; PC2188; A32266  
R/Sastry, K.; Herman, G.A.; Day, L.; Deignan, E.; Bruns, G.; Morton, C.C.; Ezekowitz, R. J. Exp. Med. 170, 1175-1189, 1989  
A/Title: The human mannose-binding protein gene. Exon structure reveals its evolutionary  
A/Reference number: J01115; MUID:90010778; PMID:2477486  
A/Accession: J01115  
A/Molecule type: DNA  
A/Residues: 1-248 <SAS>  
A/Cross-references: UNIPROT:P11226; EMBL:X15422; NID:G34486; PIDN:CAA33462.1; PID:G34487  
R/Taylor, M.E.; Brickell, P.M.; Craig, R.K.; Summerfield, J.A.  
Biochem. J. 262, 763-771, 1989  
A/Title: Structure and evolutionary origin of the gene encoding a human serum mannose-binding  
A/Reference number: S05641; MUID:90073571; PMID:2590164  
A/Accession: S05641  
A/Molecule type: DNA  
A/Residues: 1-248 <TAY>  
A/Cross-references: EMBL:X15954; NID:G34480; PIDN:CAA34079.1; PID:G1212951  
A/Accession: A34978  
A/Molecule type: protein  
A/Residues: 'X', 22-24, 'X', 26, 'X', 28-31, 'X', 33-34, 'X', 36, 'XXXX', 41-50 <TAY2>  
R/Ezekowitz, R.A.B.; Day, L.E.; Herman, G.A.  
J. Exp. Med. 167, 1034-1046, 1988  
A/Title: A human mannose-binding protein is an acute-phase reactant that shares sequence  
A/Reference number: J00027; MUID:88171281; PMID:2450948  
A/Accession: J00027  
A/Molecule type: mRNA  
A/Residues: 1-2, 'C', 4, 'IT', 8, 'S', 10-57, 'R', 59-60, 'GT', 63-106, 'PGLRRK', 113, 'SSANRNGTYQ', 115  
R/Kurata, H.; Sannoh, T.; Kozutsumi, Y.; Yokota, Y.; Kawasaki, T.  
J. Biochem. 115, 1148-1154, 1994  
A/Title: Structure and function of mannan-binding proteins isolated from human liver and  
A/Reference number: JX0319; MUID:95073978; PMID:7982896  
A/Accession: JX0319  
A/Molecule type: protein  
A/Residues: 1-248 <KUR1>  
A/Accession: PC2188  
A/Molecule type: mRNA  
A/Residues: 1-20 <KUR2>  
A/Experimental source: liver and serum  
C/Comment: Mannose-binding lectins are opsonins that are important in host defense against

C;Comment: This protein is a Ca2+-requiring animal lectin specific for mannose and N-acetylglucosamine.  
C;Gene: GDB:MBL  
A;Cross-references: GDB:120167; OMIM:154545  
A;Map position: 10q11.2-10q11.2  
A;Introns: 63/1; 102/1; 125/1  
C;Superfamily: mannose-binding lectin; C-type lectin homology  
C;Keywords: acute phase; calcium binding; endoplasmic reticulum; Golgi apparatus; hydroxymethylglutaryl-CoA lyase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-248/Product: mannose-binding lectin #status experimental <MAT>  
F;42-99/Region: collagen-like  
F;128-244/Domain: C-type lectin homology <LCH>  
F;47,73,79,82,88/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

Query Match 19.3%; Score 185.5; DB 1; Length 248;  
Best Local Similarity 30.2%; Pred. No. 4.4e-09;  
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;

QY 42 EQQALQTVCLK-----GTVHMKCFLAFTQTKTFHEASEDCISRGCTLSTPTQSGE 92  
DB 111 ERKALQTEMARIKKWLTFSGLKQVGNKFFLTNGEIMTFKVKALCVKQASVATPRNAE 170  
QY 93 NDALYELVRSQVNEAEIWLGLNDMAAGTVDWMTGARIYKQWETETIAQP--DGKTEN 151  
DB 171 NGAIQNLKE-----EALGIDTEKTEGQVLDLTGNRLTYTNW---EGEPNAGSD 221  
QY 152 CAVLSGAANGKWFDKRCRDLPYICQFIV 180  
DB 222 CVLL--LKNQOWNDVPCNSTSHLAVCFPI 248

RESULT 8  
S52781  
neurocan - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C;Accession: S52781  
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.  
submitted to the EMBL Data Library, February 1995  
A;Description: Amino acid sequence of mouse neurocan and brevican and their different exons.  
A;Reference number: S52781  
A;Accession: S52781  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1268 <RAU>  
C;Cross-references: UNIPROT:P55066; EMBL:X84727; NID:g758629; PIDN:CAA59216.1; PID:g758629  
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF  
F;176-253/Domain: link protein repeat homology <LNK1>  
F;274-355/Domain: link protein repeat homology <LNK2>  
F;964-995/Domain: EGF homology <EGF>  
F;1040-1160/Domain: C-type lectin homology <LCH>  
F;1167-1223/Domain: complement factor H repeat homology <FHD>

Query Match 19.0%; Score 182.5; DB 2; Length 1268;  
Best Local Similarity 31.8%; Pred. No. 5.4e-08;  
Matches 47; Conservative 22; Mismatches 52; Indels 27; Gaps 8;

QY 48 TVCLKGT-----KVHMKCFLAFTQTKTFHEASEDCISRGCTLSTPTQSGENDALY 99  
DB 1031 SLCEKDEGCRGWHKFGCHCYRYFAHRAWEADAERDCRRAGHLTSVHSPPEHKFI--- 1087  
QY 100 LRQSVGNEAEIWLGLNDMAAG--TWDMTGARIYKQWETETIAQPD-----GKTENCA 153  
DB 1088 --NSFGHE--NSWIGLNDRTVERDFQWNTDNG--LQYENWREK---QPDNFFAGG--EDCV 1137  
QY 154 VLSGAANGKWFDKRCRDLPYICQFIV 181  
DB 1138 VVVAHESGRWVNDVPCNLYPYCKGTV 1165

RESULT 9  
S57653  
brevican precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S57653  
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.  
submitted to the EMBL Data Library, May 1995  
A;Description: Cloning and sequence of mouse neurocan and brevican and their different exons.  
A;Reference number: S57653  
A;Accession: S57653  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-883 <RAU>  
C;Cross-references: UNIPROT:O61361; EMBL:X87096; NID:g886889; PIDN:CAA60575.1; PID:g886889  
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF  
F;49-138/Domain: immunoglobulin homology <IMM>  
F;173-250/Domain: link protein repeat homology <LNK1>  
F;271-352/Domain: link protein repeat homology <LNK2>  
F;626-657/Domain: EGF homology <EGF>  
F;664-784/Domain: C-type lectin homology <LCH>  
F;791-847/Domain: complement factor H repeat homology <FHD>

Query Match 18.8%; Score 180; DB 2; Length 883;  
Best Local Similarity 28.0%; Pred. No. 5.9e-08;  
Matches 45; Conservative 26; Mismatches 62; Indels 28; Gaps 5;

QY 40 LKEQALQTVCLK-----TKVHM-----KCFLAFTQTKTFHEASEDCIS 82  
DB 638 LEEKGFRLCLPGYGGDLCDVGLHFCSPGWEAFQGCYKHFSTRSWEAEASQCRALGA 697  
QY 83 TLSTPTQSGENDALYELVRSQVNEAEIWLGLNDMAAGTVDWMTGARIYKQWETETIA 142  
DB 698 HLTSICTPEEQDFVNDVREYQ-----WIGLNDRTIEGDFLWSDGAPLLYNNW---PG 748

RESULT 10  
JN0450  
conglutinin precursor - bovine  
N;Alternate names: C3b-binding protein  
N;Contains: conglutinin-N  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
R;Suzuki, Y.; Yin, Y.P.; Makino, M.; Kurimura, T.; Wakamiya, N.  
Biochem. Biophys. Res. Commun. 191, 335-342, 1993  
A;Title: Cloning and sequencing of a cDNA coding for bovine conglutinin.  
A;Reference number: JN0450; MUID:93213261; PMID:8460993  
A;Accession: JN0450  
A;Molecule type: mRNA  
A;Residues: 1-371 <SUZ>  
C;Cross-references: UNIPROT:P23805; DBJ:D14085; NID:g285643; PIDN:BAA03170.1; PID:g285643  
A;Experimental source: liver  
R;Kawasaki, N.; Itoh, N.; Kawasaki, T.  
Biochem. Biophys. Res. Commun. 198, 597-604, 1994  
A;Title: Gene organization and 5'-flanking region sequence of conglutinin: A C-type mammary  
A;Reference number: JN0450; MUID:94128104; PMID:8297370  
A;Accession: JN0450  
A;Molecule type: mRNA  
A;Residues: 1-371 <KA2>  
A;Note: The authors translated the codon GAT for residues 250 and 270 as Glu  
R;Lu, J.; Laurson, S.B.; Thiel, S.; Jensenius, J.C.; Reid, K.B.M.  
Biochem. J. 292, 157-162, 1993  
A;Title: The cDNA cloning of conglutinin and identification of liver as a primary site of  
A;Reference number: S33235; MUID:93277452; PMID:7684896  
A;Accession: S33235  
A;Molecule type: mRNA  
A;Residues: 1-172, 'H', 174-217, 'A', 219-271, 'V', 273-371 <LUJ>  
C;Cross-references: EMBL:X71774; NID:g395267; PIDN:CAA50665.1; PID:g395268  
A;Experimental source: liver  
R;Lee, Y.M.; Leiby, K.R.; Allar, J.; Paris, K.; Lerch, B.; Okarna, T.B.  
J. Biol. Chem. 266, 2715-2723, 1991









Db	228	QOMEALNGKQL--EAFSRYKKA--ALPPDQSVGDKIFRAANSEPFEDAKEMCROA	283
Qy	81	GGTLSTPQTGSNDALYELROSQVGNBARIWGLNDMAAEGTWDMTGARIAYKNWETEI	140
Db	284	GGQLASPRSATENAQQQLV---TAHSKAAFLSMTDVGTEGKFTYPTGEALVYSNW--A	337
Qy	141	TAQP-DGGKTENC AVLSGAANGKWFDRCRDQLPYICQF	178
Db	338	PGEPNNGGAENCVEI--FTNGQWNDKACGEORLVICEF	374

Search completed: April 26, 2005, 11:53:57  
Job time : 40 secs

**THIS PAGE BLANK (USPTO)**